



# **NASA Strategic Planning Document: A Comprehensive Plan for the Long-Term Calibration and Validation of Oceanic Biogeochemical Satellite Data**

*Stanford B. Hooker, Charles R. McClain, and Antonio Mannino*

## The NASA STI Program Office ... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

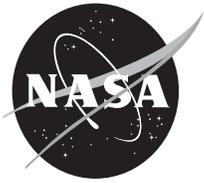
- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and mission, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at <http://www.sti.nasa.gov/STI-homepage.html>
- E-mail your question via the Internet to [help@sti.nasa.gov](mailto:help@sti.nasa.gov)
- Fax your question to the NASA Access Help Desk at (301) 621-0134
- Telephone the NASA Access Help Desk at (301) 621-0390
- Write to:  
NASA Access Help Desk  
NASA Center for Aerospace Information  
7115 Standard Drive  
Hanover, MD 21076-1320



# **A Comprehensive Plan for the Long-Term Calibration and Validation of Oceanic Biogeochemical Satellite Data**

*Stanford B. Hooker, Charles R. McClain, and Antonio Mannino  
NASA/Goddard Space Flight Center, Greenbelt, Maryland*

**National Aeronautics and  
Space Administration**

Goddard Space Flight Center  
Greenbelt, Maryland 20771

---

Available from:

NASA Center for AeroSpace Information  
7115 Standard Drive  
Hanover, MD 21076-1320  
Price Code: A17

National Technical Information Service  
5285 Port Royal Road  
Springfield, VA 22161  
Price Code: A10

## Table of Contents

ABSTRACT .....	1
1. Introduction .....	1
1.1 Calibration and Validation .....	2
1.2 Project Offices .....	3
2. Lessons Learned .....	4
2.1 Publish Protocols .....	5
2.2 Estimate Uncertainties .....	6
2.3 Establish Performance Metrics .....	8
2.4 Provide Access to High-Quality Data .....	9
2.5 Manage Vicarious Calibration Sites .....	10
2.6 Address Optical Complexity .....	10
2.7 Develop and Evaluate Instrumentation .....	11
3. The Proposed Activity .....	12
3.1 Satellite Data Processing .....	15
3.2 Calibration and Validation .....	16
3.3 Competed Elements .....	17
4. Issues and Discussion .....	18
4.1 Accountability .....	18
4.2 Above- versus In-Water Radiometry .....	20
4.3 Hyperspectral Radiometry .....	20
4.4 Ultraviolet Wavelengths .....	21
4.5 Primary Productivity .....	21
4.6 The Advanced Science Plan .....	22
5. Strategic Summary .....	23
ACKNOWLEDGMENTS .....	26
GLOSSARY .....	26
SYMBOLS .....	27
REFERENCES .....	27



## ABSTRACT

The primary objective of this planning document is to establish a long-term capability for calibrating and validating oceanic biogeochemical satellite data. It is a pragmatic solution to a practical problem based primarily on the lessons learned from prior satellite missions. All of the plan's elements are seen to be interdependent, so a horizontal organizational scheme is anticipated wherein the overall leadership comes from the NASA Ocean Biology and Biogeochemistry (OBB) Program Manager and the entire enterprise is split into two components of equal stature: calibration and validation plus satellite data processing. The detailed elements of the activity are based on the basic tasks of the two main components plus the current objectives of the Carbon Cycle and Ecosystems Roadmap. The former is distinguished by an *internal* core set of responsibilities and the latter is facilitated through an *external* connecting-core ring of competed or contracted activities. The core elements for the calibration and validation component include a) publish protocols and performance metrics; b) verify uncertainty budgets; c) manage the development and evaluation of instrumentation; and d) coordinate international partnerships. The core elements for the satellite data processing component are e) process and reprocess multisensor data; f) acquire, distribute, and archive data products; and g) implement new data products. Both components have shared responsibilities for initializing and temporally monitoring satellite calibration. Connecting-core elements include (but are not restricted to) atmospheric correction and characterization, standards and traceability, instrument and analysis round robins, field campaigns and vicarious calibration sites, *in situ* database, bio-optical algorithm (and product) validation, satellite characterization and vicarious calibration, and image processing software. The plan also includes an accountability process, creating a Calibration and Validation Team (to help manage the activity), and a discussion of issues associated with the plan's scientific focus.

## 1. Introduction

The global mapping of the oceanic biosphere is accomplished through the determination of radiometric quantities. Specifically, the values of the spectral radiances at the top of the atmosphere, from which (after atmospheric correction), the spectral radiances emerging from the ocean surface,  $L_W(\lambda)$ , are extracted ( $\lambda$  denotes wavelength). These so-called *water-leaving radiances* are a critical part of the success of an ocean color—or alternatively, ocean reflectance—satellite mission, which is determined by the quality of the remote sensing data and the availability of the derived products. The former is provided by a calibration and validation paradigm, and the latter by a data processing capability. Both components require several important activities, discussed in more detail below, and the need to achieve an agreed upon accuracy requires cooperation between the organizational elements.

Because of the focus on satellite observations, the ultimate success and future expansion of the OBB Program is inexorably tied to launching new missions based on novel research topics and assuring the quality of the ensuing satellite data. Both of these objectives require effective interactions between the scientific research community and the calibration and validation activity. The plan espoused here is based on more than just synergism—the goal is to integrate the two work areas into a single enterprise.

The long-term OBB programmatic requirements are articulated in an Advanced Science Plan, *On the Shores of a Living Ocean: The Unseen World*, which was drafted by

a subset of the scientific community led by the OBB Program Manager†. The designated mission themes from this plan, along with the corresponding high-priority research questions, highlight the science and mission concepts the calibration and validation activity must help enable.

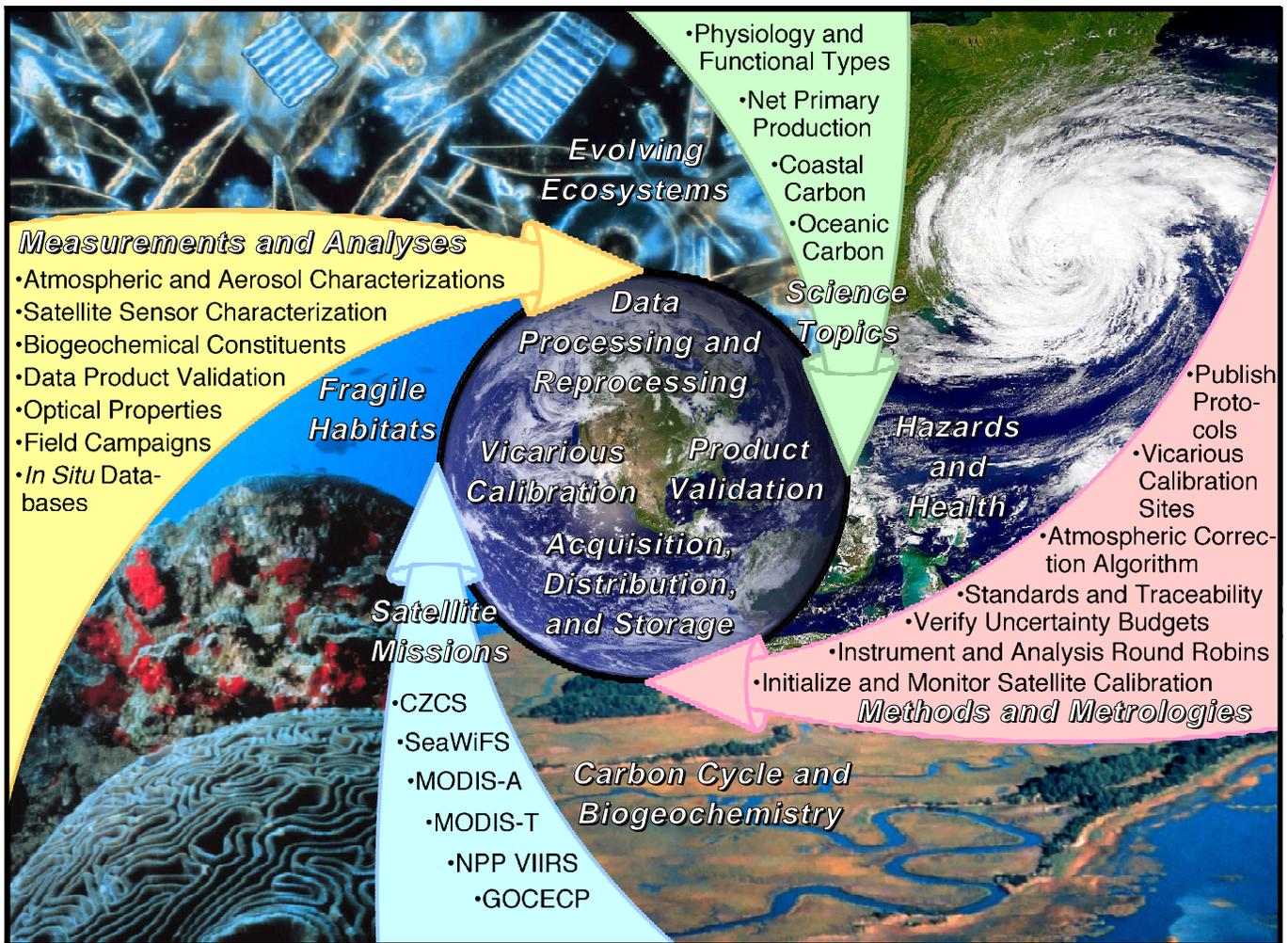
The mission themes span a range of scales and applications: a) global separation of pigments and ecosystem components, b) high spatial and temporal resolution of coastal waters, c) active assessment of plant physiology and composition, and d) determination of mixed layer depths. The corresponding research questions span equally large scales:

- How are oceanic ecosystems and their attendant biodiversity influenced by climate or environmental changes, and how will these evolve over time?
- How do carbon and other elements transition between oceanic pools and pass through the Earth system, and how do biogeochemical fluxes impact the ocean and planetary climate over time?
- How (and why) are the diversity and geographical distribution of coastal marine habitats changing, and what are the implications for human health?
- How do hazards and pollutants impact the hydrography and biology of the coastal zone and human activities, and can the effects be mitigated?

The successful implementation of the science and mission concepts inevitably leads to technology development

---

† The Advanced Science Plan is available from the following Web site: <http://oceancolor.gsfc.nasa.gov/DOCS>.



**Fig. 1.** An integrated depiction of the four scientific research areas (with pictorial representations), the primary elements of the global calibration and validation activity (center), and the principal programmatic responsibilities divided into four categories (colored arrows). The satellite missions are discussed in Sect. 1.2.

issues, which further impact the measurements and analyses to be performed, as well as the methods and metrologies to be used. Although this increasing level of detail makes it difficult to summarize the needed functions, requirements, and objectives in a solitary depiction, it is useful to do so, because all must be fulfilled within a single structure—the OBB Program. An integrated perspective of the main research areas, the primary calibration and validation elements, and the principal programmatic responsibilities are shown in Fig. 1. These three partitions establish the scale and complexity of what must be undertaken. The remainder of this document is devoted to how the calibration and validation component can be accomplished at the requisite quality to support the other two.

## 1.1 Calibration and Validation

In remote sensing applications, “calibration” and “validation” can have alternative meanings to different individuals and communities. Some think of the two as being

distinct and separate activities, while others view them as tightly connected and interdependent. Calibration is frequently defined as the prelaunch characterization followed by the continuing analysis of the onboard sensor calibrators once on-orbit operations commence. Validation is usually thought of as the development of data processing schemes (e.g., atmospheric correction and derived geophysical quantities), plus the verification of product accuracies using ground-truth data. It is not unusual for these elements to be considered part of the same function.

For the purposes of this document, “calibration” is associated with those activities needed to ensure a proper prelaunch characterization of the satellite sensor, tracking the postlaunch sensor performance over time, plus the vicarious† adjustment of the sensor’s prelaunch calibration

† In this context, “vicarious” simply admits that the preferred rigor of actually calibrating a satellite sensor on orbit is not a practical possibility, so a substitute—but agreeably robust—procedure is being used instead.

to match high-quality *in situ* observations. The “validation” component consists of the myriad tasks required to establish the efficacy of the data products derived from an *algorithm*† applied to the observations recorded by a properly calibrated sensor. A more compact vocabulary is simply to refer to this entire enterprise as “vicarious calibration and product validation,” and, ultimately, as the even more succinct (and originally ambiguous) “calibration and validation.”

The overlap between calibration and validation occurs because both activities require ground-truth—more properly sea-truth—observations. Calibration requires greater accuracy than validation, so applying data from the latter to the former is usually not considered. The Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Project, for example, requires a radiometric accuracy to within 5% absolute and 1% relative, and chlorophyll *a* (Chl *a*) concentration‡ to within 35% over a range of 0.05–50.0 mg m<sup>-3</sup> (Hooker and Esaias 1993).

The difficulty in using validation data for calibration exercises is not associated with less rigorous techniques being used for validation measurements; it is simply a consequence of the dynamic range of the two activities—in fact, the same protocols are used in both cases. Vicarious calibration requires a sampling site wherein the contribution of natural variability—atmospheric and oceanic—is minimized, so the total uncertainty is properly reduced. This is most simply (but not exclusively) satisfied at a site with predominantly clear properties (skies and waters) with a simplistic particle distribution (exclusively marine aerosols for the atmosphere and in-water properties that depend primarily on Chl *a*). Such a small range in parameters represents almost a single point in the global expression of a data product, so validation requires multiple sites wherein the associated natural variability will presumably degrade the uncertainty budget required for calibration.

## 1.2 Project Offices

Whether or not calibration and validation are intertwined or separated can also depend on the complexity of the mission. For SeaWiFS, the activities are integrated into a single function closely coordinated with the data processing group (made possible because all the elements are collocated). In the Moderate Resolution Imaging Spectro-

radiometer (MODIS) program§, which reflects the original Earth Observing System (EOS) paradigm, sensor calibration is handled by one group, the MODIS Characterization Support Team (MCST), and product validation is the responsibility of the (land, ocean, and atmosphere) science teams (which may not have a close relationship with the MCST and might use their own vicarious calibrations).

The MODIS program and the structure of the MODIS ocean team is very similar to the NIMBUS-7 Experiment Team (NET) for the Coastal Zone Color Scanner (CZCS). Recently, the NASA Earth Science Program has adopted a more centralized (SeaWiFS) approach under the *missions-to-measurements* strategy, wherein, a single group handles the data processing plus many of the calibration and validation functions. Ocean biogeochemistry was the first discipline to adopt this model with the GSFC Ocean Biology Processing Group assuming that role.

Aside from any organizational options, the strategies for executing the underlying tasks have also evolved. For the purposes of this document, a brief review of the relevant programs is appropriate, because this helps defend the strengths of the recommended approach. In addition, the Advanced Science Plan is embracing a broader set of science objectives than before, which must be reflected in the calibration and validation strategy.

The CZCS was a proof-of-concept mission with modest science and data processing goals that were brought to fruition—and greatly exceeded—during the 1980s. Field campaigns for algorithm development were conducted prior to launch in 1978, and postlaunch validation experiments were concentrated in the first year of operations (Gordon et al. 1980 and 1983). The main problem was characterizing the degradation of the visible bands without a monitoring capability. Evans and Gordon (1994) evaluated the decay by assuming constant clear-water radiances, but the lesson was that a robust calibration program spanning the duration of the mission would be needed in the future.

The SeaWiFS and MODIS missions, designed with the CZCS experience in mind, were developed in parallel and leveraged a number of joint developments, e.g., the atmospheric correction scheme and the Marine Optical Buoy (MOBY) vicarious calibration site (Clark et al. 1997). As noted above, the organizational and financial structures were very different. The SeaWiFS calibration and validation activity (McClain et al. 1992) had a well defined budget with considerable flexibility in apportioning funds between internal and external components. As a result, the documentation of field protocols, the development of new instruments (e.g., the SeaWiFS Transfer Radiometer and SeaWiFS Quality Monitor), plus the calibration, pigment, and data analysis round robins, were directly supported (Hooker and McClain 2000, and McClain et al. 2004).

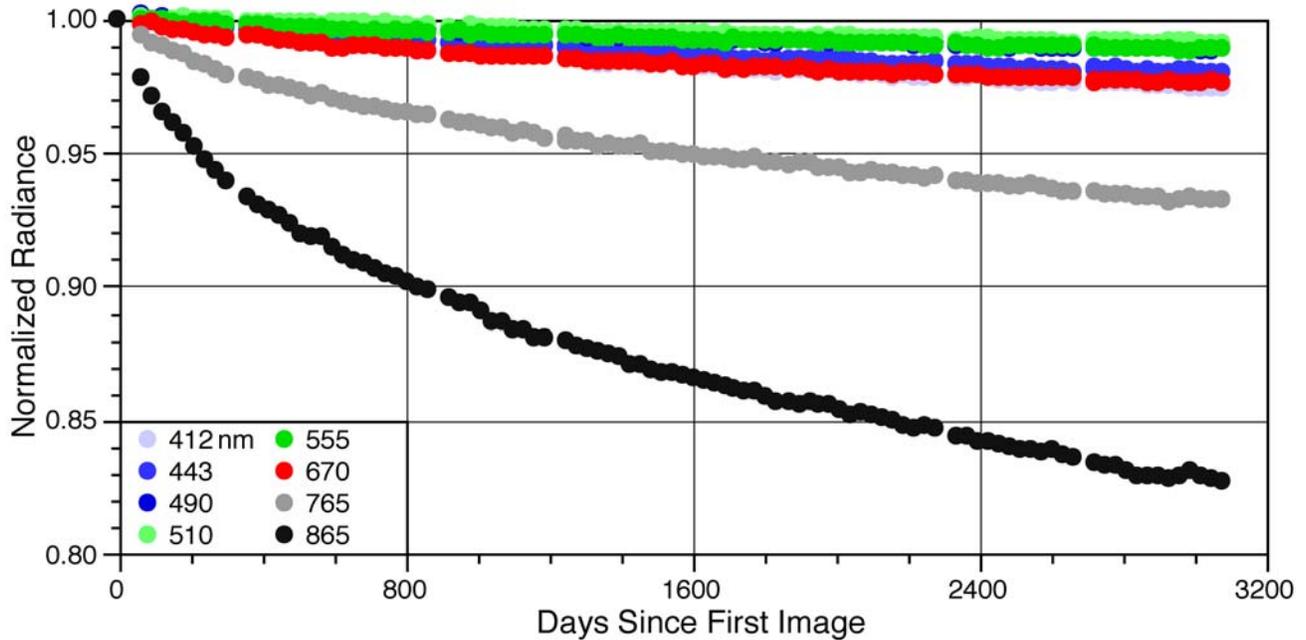
---

† “Model” would be a more appropriate term, but the simplicity of most of the relationships involved—parameter *y* is obtained directly from observation *x* using a straightforward and easily implemented mathematical equation (e.g., the derivation of the chlorophyll *a* concentration from reflectance ratios with a polynomial function)—makes “algorithm” a widely accepted choice.

‡ In fact, field-to-satellite comparisons (or matchups) are made with respect to the total chlorophyll *a* (TChl *a*) concentration, denoted [TChl *a*].

---

§ There are two MODIS instruments, which were launched on the Terra and Aqua satellites, and are denoted MODIS-T and MODIS-A, respectively.



**Fig. 2.** The normalized radiances (with respect to the first measurement) derived from SeaWiFS images of the full Moon. The colored circles correspond to the generalized wavelengths: blue for 412, 443, and 490 nm; green for 510 and 555 nm; red for 670 nm; and black for the near-infrared (NIR) bands 765 and 865 nm.

SeaWiFS did not have a formal instrument team like the MODIS ocean team (Esaias et al. 1998), and relied heavily on the latter while supporting additional field work such as the Atlantic Meridional Transect (AMT) cruises (Aiken et al. 2000). Both SeaWiFS and MODIS use onboard calibration techniques to track sensor stability—capabilities CZCS did not have, but were outlined by Gordon (1987), and MOBY is the only data source for adjusting the calibration gains after the time dependencies are removed (Barnes et al. 2001 and Eplee et al. 2001).

The Sensor Intercomparison and Merger for Biological and Interdisciplinary Oceanic Studies (SIMBIOS) activity was not a flight project, but its goal was the intercalibration and product validation of ocean color sensors (McClain and Fargion 1999, and McClain et al. 2002). The organizational structure was a SeaWiFS and MODIS hybrid, in the sense that it had a project manager and a collocated data analysis and processing group, but it also had a science team very similar to the MODIS ocean team (the latter was included in the SIMBIOS science team). SIMBIOS assumed the responsibility of continuing and expanding a number of activities initiated by the SeaWiFS program at a time when the SeaWiFS budget was ramping down after launch (as originally planned).

SeaWiFS calibration and validation currently continues only at a level needed to support the lunar and solar analyses, as well as occasional reprocessings to keep the atmospheric correction and bio-optical algorithms up to date. The MODIS ocean team was recomputed in 2003 and is expected to continue throughout the Aqua and Terra missions. SeaWiFS and MODIS were originally envisioned to

be primary elements of an international effort to develop a long-term time series of global satellite observations (Abbott et al. 1994), an objective that has been realized. Although continuous observations were a stated priority in the early 1990s, much remained to be learned about producing a *climate data record* (CDR), and its importance to science-quality research (McClain et al. 2006).

Maintaining a CDR time series is a continuing goal and will depend on whether the National Polar Orbiting Environmental Satellite System (NPOESS) Visible and Infrared Imaging Radiometer Suite (VIIRS) delivers high-quality data, because no NASA ocean color mission after MODIS is approved. Nonetheless, mission concepts continue to be developed and provide needed insights into the capabilities and requirements for next-generation spaceborne sensors, for example, the Global Ocean Carbon, Ecosystems, and Coastal Processes (GOCECP) mission.

## 2. Lessons Learned

A primary CZCS lesson was that accurately tracking sensor stability over the course of a mission is essential. Consequently, the SeaWiFS design included solar and lunar calibration gains, a solar diffuser, and a strategy for monthly images of the Moon at approximately a  $7^\circ$  phase angle, whereas MODIS incorporated a solar diffuser with a stability monitor. Both approaches have proven to be robust, although, years after launch, the methodologies continue to be refined. Figure 2 presents the SeaWiFS lunar time series and shows the degradation in SeaWiFS varies smoothly over time and occurs primarily in the NIR bands

(unlike CZCS). MODIS-T has experienced sudden sensitivity shifts related to the spacecraft electronics, which were resolved in the solar calibration data, but would not have been adequately captured in monthly lunar observations. The MOBY match-up time series is inadequate for tracking either the SeaWiFS or the MODIS-T degradations. Future missions, therefore, require robust onboard measurement capabilities and strategies for tracking sensor stability. Considerable effort is being made to ensure VIIRS doesn't, but the inability to make lunar observations means unpredictable degradations (like MODIS-T experienced) will be difficult to detect and characterize.

Aside from onboard stability tracking, prelaunch sensor characterization is critical, because ocean color data products are sensitive to 0.1% calibration uncertainties. The system-level response of the instrument to top-of-the-atmosphere radiances is a function of many design attributes, which are measured during prelaunch tests over an appropriate range of parameter variations (e.g., temperature and scan geometry). The resulting functional relationships are convolved into the overall calibration equation relating volts to radiance. Sensor responsivity versus illumination using a calibration source (Johnson et al. 1999a), is one of many such tests. Prelaunch characterization uncertainties translate into unwanted variations in the derived radiances<sup>†</sup>. Usually, the parameters used with the functional relationships cannot be derived on orbit, so a significant effort is required to evaluate instrument design, prelaunch characterization, and postlaunch stability.

Assuming the requisite expertise for prelaunch characterization and on-orbit monitoring are available, the final components of a comprehensive plan for ocean color calibration and validation is undertaken here by first admitting that the final requirements are practical ones:

*High-quality data are needed for both vicarious calibration and product validation exercises. The data must be produced using approved sampling, analysis, and reporting protocols, wherein the resulting quality assurance (QA) parameters are shown to be within community-established performance metrics (e.g., accuracy thresholds). The data must also be collected across the requisite dynamic range associated with the primary variables.*

The last requirement is associated primarily with validation exercises, because the basic concept associated with

vicarious calibration is to minimize the influence of as many natural sources of variance as possible.

Regardless of differences in perspective, calibration and validation require match-up data, that is, contemporaneous observations by the satellite and an *in situ* instrument. In most cases, variables that explicitly account for the solar irradiance,  $E_d(0^+, \lambda)$ , at the time of data collection—so-called apparent optical properties (AOPs)—are used for match-up analysis, e.g., the remote sensing reflectance,  $R_{rs}(\lambda)$ , the radiance reflectance,  $\rho_w$ , or the normalized water-leaving radiance,  $[L_W(\lambda)]_N$ . This normalization<sup>§</sup> by the illumination conditions makes  $R_{rs}$  the primary variable for estimating chlorophyll *a* concentration from *in situ* optical measurements (O'Reilly et al. 1998), which means it is a central variable for validation exercises. For vicarious calibration, a final—more exact—computation includes correcting the observations for the angular (bidirectional) dependence of  $L_W$  (Mueller and Morel 2003), which are also used for validation and routine data processing.

The next step in the pragmatic approach adopted here is to discuss the particular elements—or, more properly, objectives—which are critical to the execution of a plan devised to achieve the above requirements. This discussion is formulated in terms of present and past calibration and validation capabilities. Fundamentally, this means reviewing the lessons learned from the paradigms discussed earlier (CZCS through MODIS), and then making sure successful procedures are part of the plan, and any needed corrections or additions are properly identified and incorporated.

## 2.1 Publish Protocols

To ensure the needed field measurements were in keeping with the remote sensing requirements, the SeaWiFS Project convened a workshop to draft the SeaWiFS Ocean Optics Protocols (hereafter referred to as the Protocols). The Protocols initially adhered to the Joint Global Ocean Flux Study (JGOFS) sampling procedures (JGOFS 1991) and defined the standards for optical measurements to be used in SeaWiFS calibration and validation activities (Mueller and Austin 1992). Over time, the Protocols were revised (Mueller and Austin 1995), and then recurrently updated essentially on an annual basis over the duration of the SIMBIOS Project to include a full suite of biogeochemical parameters (Mueller 2000, 2002, and 2003).

<sup>†</sup> At the insistence of the ocean color members of the NPOESS Preparatory Project Science Team, the solar diffuser is being redesigned to minimize earthshine and other sources of illumination contamination of the solar diffuser. The sensitivity for this problem is a result of recognized, but uncorrected, earthshine in both MODIS diffuser measurements.

<sup>‡</sup> Barnes et al. (1994a, 1994b, 1995) provide excellent examples of characterization testing and data analysis (e.g., polarization, stray light, point spread response, and solar diffuser bidirectional reflectance).

<sup>§</sup> Derivations of  $L_W(\lambda)$  in identical waters, but different illumination conditions, will differ. The variability can be removed, in part, by normalizing  $L_W(\lambda)$  by the solar irradiance to compute  $R_{rs}(\lambda) = L_W(\lambda)/E_d(0^+, \lambda)$  or  $\rho_w = \pi R_{rs}(\lambda)$ . Another appropriate choice is to use  $[L_W(\lambda)]_N$ , which is defined as the hypothetical water-leaving radiance that would be measured with no atmospheric loss and a zenith sun at the mean Earth–Sun distance (Gordon and Clark 1981). The latter requires an adjustment to  $R_{rs}(\lambda)$  by the time-dependent mean extraterrestrial solar irradiance,  $F_0(\lambda, d)$ , which is usually formulated to depend on the sequential day of the year,  $d$ .

The Protocols represent a unique accomplishment, and a significant lesson they confirm is that the state of the art is advanced by quantifying methodological uncertainties. An uncertainty analysis can only commence if the requisite procedures can be accurately implemented, which means they must be properly documented. Once this information is available, a next-generation capability can be measured against the current one to determine whether or not progress is being made. In addition, an uncertainty analysis can show how much of the reported variance is real (and, thus, mostly unavoidable) and how much is artificial (usually removable—or at least reducible).

The utility of a set of Protocols that are endorsed and maintained by a broader community, therefore, far exceeds the simple accomplishment of providing the procedures for accomplishing certain tasks or measurements. As long as they are a *work in progress*, updating the Protocols provides a periodic review of the state of the art and gives new ideas or procedures a forum for evaluation. This opportunity to discuss and document how the basic tools for meeting calibration and validation requirements are being satisfied is a critical element of a successful program.

## 2.2 Estimate Uncertainties

To maintain internal consistency between calibrations of *in situ* radiometers and the satellite sensor, the SeaWiFS Project (as part of the Protocols) required traceability of calibration standards to the National Institute of Standards and Technology (NIST), which is now a requirement for all domestic (ocean color) satellite missions. The Project also implemented a series of SeaWiFS Intercalibration Round-Robin Experiments (SIRREXs) to investigate and minimize uncertainties associated with AOP instruments, because the SeaWiFS sea-truth uncertainty budget can only be satisfied if each contributing uncertainty is on the order of 1–2% (Hooker and McClain 2000). As a generalized description, this constitutes so-called *1% radiometry*; in other words, uncertainty sources in the calibrated use of a sensor must be kept at about the 1% level.

In the progression from the first to the third SIRREX (Mueller 1993, Mueller et al. 1994, and Mueller et al. 1996, respectively), uncertainties in the traceability to NIST for intercomparisons of spectral lamp irradiance and sphere radiance improved from 7–8% to 1–2%, respectively. The fourth through seventh SIRREX activities further investigated laboratory and field protocols (Johnson et al. 1996, Johnson et al. 1999b, Riley and Bailey 1998, and Hooker et al. 2002, respectively), and showed calibration uncertainties of about 2–3% were routinely achievable if the Protocols were carefully executed†. More recently, SIRREX-8 revealed the immersion factors supplied by a commercial

manufacturer were more than 10% in error at some wavelengths (Zibordi et al. 2002a), and there are other examples of the need for independent instrumentation evaluations (e.g., Mueller 1995 and Hooker and Maritorena 2000).

The uncertainties associated with data processing are tied to the original instrument characterizations and sampling protocols, but there are subjective aspects that are not completely resolved by a single protocol. The first SeaWiFS Data Analysis Round Robin (DARR-94) investigated the uncertainties in the data processing of in-water optical profiles and showed differences in commonly used methods for determining primary optical parameters were about 3–4% of the aggregate mean estimate (Siegel et al. 1995). The focus of the second DARR (DARR-00) was to determine if these results could be improved upon (Hooker et al. 2001). In terms of overall spectral averages, many of the DARR-00 intercomparisons were to within 2.5%, and if the processing options were made as similar as possible, agreement to within less than 1% was routinely possible for two of the processors. Much higher uncertainties were documented, however, and many of these were associated with data products critical to calibration and validation.

Optical parameters do not account for all of the validation requirements. The proper determination of [TChl *a*] is central to the objectives of all ocean color missions. The SeaWiFS High Performance Liquid Chromatography (HPLC) Analysis Round-Robin Experiment (SeaHARRE) investigated the uncertainties in the quantitation of marine pigments (Hooker et al. 2000a and Hooker et al. 2005). The SeaHARRE-1 results (Claustre et al. 2004) showed the determination of [TChl *a*] was to within 8% (well within remote sensing requirements‡), whereas the quantitation of the common carotenoids was less accurate and on the order of 24%. The average SeaHARRE-2 [TChl *a*] uncertainty was 11.4%, but only 7.8% for a quality-assured subset of four methods (denoted *A'*). Using a QA procedure based on a limit of quantitation (LOQ) threshold and choosing the *A'* subset as the proxy (or reference) for truth in the uncertainty calculations, reduced the average [TChl *a*] uncertainty to 5.9% (and 17.2% for the other laboratories). Applying an LOQ threshold to the SeaHARRE-1 data resulted in a similar uncertainty in [TChl *a*] of 5.5%.

The recurring (essentially annual) inquiries into uncertainties establish an increasingly detailed calibration and validation uncertainty budget, which is presented in Table 1. The entries show the difficulty of maintaining the aforementioned radiometric 1% uncertainty requirements, as well as the ensuing increase in variance when data from a diverse set of contributors are used for algorithm development or validation activities. This is an important point,

‡ Based on an agreement to within 35% (the SeaWiFS requirement) and assuming the sources of uncertainty combine independently (i.e., in quadrature), an upper accuracy range of about 25% is probably acceptable,  $\sqrt{35^2/2}$ , although 15% would presumably permit significant algorithm refinement.

† The SIMBIOS Radiometric Intercomparison (SIMRIC) activity largely confirmed this level of achievement (Meister et al. 2002 and 2003).

**Table 1.** An example—and necessarily incomplete—summary of representative uncertainties (in percent) associated with ocean color calibration and validation activities as determined primarily from an AOP measurement perspective in open-ocean waters and the SIRREX, DARR, and SeaHARRE investigations, as well as a variety of field campaigns conducted during the time period of the relevant round-robin laboratory exercises. The possible uncertainties are divided into those expected of field teams working specifically for satellite missions, and those expected from the broader scientific community contributing data to large databases. The former are further divided into minimum, typical, and maximum expectations, whereas the latter are divided into overall (average) and worst-case values. The sources of uncertainties fall into six groups: a) instrument characterization, with the absolute calibration of the light sensors being the most important; b) deployment effects, some of which are correctable and others are minimizable; c) natural variability; d) data processing, with distinctions for instrument types and processors provided by the instrument manufacturer; e) intercalibrated systems, which represent a less independent and, therefore, reduced set of uncertainty sources; and f) pigment concentrations, derived from HPLC analysis.

<i>Source of Uncertainty</i>	<i>Field Teams</i>			<i>Community Database</i>	
	<i>Minimum</i>	<i>Typical</i>	<i>Maximum</i>	<i>Overall</i>	<i>Worst-Case</i>
1. Absolute Radiometric Calibration	1.5	2.7	3.5	2.9	6.3
2. Immersion Factor	0.1	0.3	1.0	2.7	10.9
3. Dark Current	0.1	0.2	0.5	0.4	1.9
4. Cosine Response	1.4	1.8	2.2	2.2	6.5
5. <i>In Situ</i> (Electro-optical) Stability	0.1	0.4	1.2	1.1	8.0
<i>Representative Uncertainty</i>	<i>1.7</i>	<i>2.9</i>	<i>4.0</i>	<i>4.3</i>	<i>15.4</i>
6. In-Water Self-Shading Correction	0.6	1.2	2.6	1.5	3.8
7. In-Water Platform Perturbations	0.0	0.5	1.2	2.7	8.2
8. Above-Water Bidirectional Correction	0.4	0.5	0.6	7.7	11.2
9. Above-Water Platform Perturbations	0.1	0.2	0.3	7.2	20.7
10. Deployment (Mechanical) Stability	0.0	0.2	1.0	0.5	2.4
<i>Representative Uncertainty</i>	<i>0.2</i>	<i>0.6</i>	<i>1.5</i>	<i>2.3</i>	<i>6.9</i>
11. Case-1 Environmental Variability	0.3	1.1	3.1	4.9	7.5
12. Coastal Environmental Variability	2.0	3.1	5.3	7.8	13.8
<i>Representative Uncertainty</i>	<i>0.9</i>	<i>1.8</i>	<i>3.8</i>	<i>5.9</i>	<i>9.6</i>
13. Winch and Crane Data Processing	0.1	0.5	1.1	3.8	27.1
14. Free-Fall Data Processing	0.1	0.3	1.0	3.5	11.7
15. Commercial Spectral Data Processing	4.8	6.2	10.7	7.0	21.6
16. Commercial Band-ratio Data Proc.	0.1	2.4	4.2	4.1	17.1
<i>Average Uncertainty</i>	<i>1.3</i>	<i>2.4</i>	<i>4.3</i>	<i>4.6</i>	<i>19.4</i>
17. In-Water Intercalibrated Method	1.5	1.9	3.3	4.3	16.5
18. Above-Water Intercalibrated Method	0.6	2.1	3.6	9.4	13.7
19. Above- and In-Water Intercal. Method	1.2	1.7	2.2	6.7	20.9
<i>Average Uncertainty</i>	<i>1.1</i>	<i>1.9</i>	<i>3.0</i>	<i>6.8</i>	<i>17.0</i>
20. C-8 HPLC [TChl <i>a</i> ]	5.0	6.3	7.9	11.4	23.8
21. C-18 HPLC [TChl <i>a</i> ]	7.1	12.9	15.2	19.5	28.9
22. Spectrophotometric [TChl <i>a</i> ]	3.9	4.9	7.1	9.8	20.1
23. Carotenoid Pigment Concentration	4.2	12.3	36.9	34.3	55.1
<i>Average [TChl <i>a</i>] Uncertainty</i>	<i>5.3</i>	<i>8.2</i>	<i>10.1</i>	<i>13.6</i>	<i>24.3</i>

*Notes:*

1. Based primarily on the calibration of upwelled radiance ( $L_u$ ) sensors.
2. A combination of radiance and irradiance values, with the former being used as typical.
4. Applicable only to irradiance sensors (but included for completeness).
7. Assumes many deployment configurations rely on winch and crane systems.
8. Representative of the uncertainty in the input parameters used to calculate these quantities and not the intrinsic uncertainties in the look-up tables being used.
12. A combination of Case-1 and Case-2 water types, with the former predominant.
15. “Commercial” refers to a data processor supplied by an instrument manufacturer (also applicable to item 16).
22. Assumes there is sufficient pigment load for the technique to be appropriate.
23. The nine carotenoids associated with the so-called primary pigments (PPig).

**Table 2.** The performance metrics for the four categories established for validating the determination of marine pigments using an HPLC method: concentration (average precision,  $\bar{\xi}$ , and accuracy,  $|\bar{\psi}|$ , for TChl *a* and the primary pigments, PPig $\dagger$ ); separation (minimum resolution,  $\check{R}_s$ , and average retention time precision,  $\bar{\xi}_{t_R}$ ); average injection precision,  $\bar{\xi}_{inj}$  (the average of an early- and late-eluting pigment standard, e.g., Perid and Chl *a*); and calibration (the average absolute percent differences of the residuals to the calibration fit for Chl *a*,  $|\bar{\psi}|_{res}$ , and the precision of the dilution devices,  $\bar{\xi}_{cal}$ ). The PPig and TChl *a* performance metrics are based on using the analysis of a mixture of laboratory standards and replicate field samples with approximately equal weights applied to each (remembering that uncertainties are assumed to combine in quadrature and that the latter presupposes the inclusion of replicate filter collection during field sampling). The corresponding values for the Horn Point Laboratory (HPL) method (Van Heukelem and Thomas 2001) are given as an example, the overall performance of which is considered “state-of-the-art,” because the average score of the weights is 3.7,  $(4 + 4 + 4 + 3 + 3 + 4 + 4 + 4 + 3 + 4)/10$ .

Performance Weight, Category, and Score	TChl <i>a</i>		PPig		Separation $\ddagger$		Injection $\S$ ( $\bar{\xi}_{inj}$ )		Calibration $\P$	
	$\bar{\xi}$	$ \bar{\psi} $	$\bar{\xi}$	$ \bar{\psi} $	$\check{R}_s$	$\bar{\xi}_{t_R}$	Perid	Chl <i>a</i>	$ \bar{\psi} _{res}$	$\bar{\xi}_{cal}$
1. Routine 0.5	8%	25%	13%	40%	0.8	0.18%	10%	6%	5%	2.5%
2. Semiquantitative 1.5	5	15	8	25	1.0	0.11	6	4	3	1.5
3. Quantitative 2.5	3	10	5	15	1.2	0.07	4	2	2	0.9
4. State-of-the-Art 3.5	$\leq 2$	$\leq 5$	$\leq 3$	$\leq 10$	$\geq 1.5$	$\leq 0.04$	$\leq 2$	$\leq 1$	$\leq 1$	$\leq 0.5$
<i>HPL Method</i>	1	5	2	12	1.2	0.02	$< 1$	$< 1$	1.1	0.4

$\dagger$  The primary pigments are total chlorophyll *a*, *b*, and *c*, plus nine carotenoids (Hooker et al. 2005).

$\ddagger$   $\check{R}_s$  is determined from a critical pair involving a primary pigment. The retention time precision entries are computed from coefficient of variation values based on sequential replicate injections of pigments identified in a mixture of pigment standards. In the absence of a diverse set of early- through late-eluting pigments, a practical alternative is to compute  $\bar{\xi}_{t_R}$  using Perid, Fuco, Diadino, Chl *a*, and  $\beta\beta$ -Car based on three sequential injections.

$\S$  The  $\bar{\xi}_{inj}$  terms are calculated from the average of replicate injections of an early- and late-eluting pigment in the same run (Perid is chosen to include the possible effects of peak assymetry, which is not presented as a separate parameter).

$\P$  The  $|\bar{\psi}|_{res}$  values are based on calibration points within the range of concentrations typical of the SeaHARRE-2 field samples. To determine this metric for an arbitrary sample set,  $|\bar{\psi}|_{res}$  is computed using those calibration points within the range of concentrations expected in the field samples to be analyzed.

because both types of work proceed most effectively when the data dispersion is natural and not artificial. In the absence of a QA parameter or a performance metric, a mixture of high quality to worst-case data are brought together with no objective way to properly separate them.

### 2.3 Establish Performance Metrics

Performance metrics are a powerful product of an investigation into uncertainties, because they have the potential of removing the burden of maintaining an overly diverse set of protocols (that have to be continually updated) or agreeing on a single protocol that satisfies the current suite of community problems. Community priorities will necessarily evolve, and at times rather rapidly, so it is perhaps appealing to be able to set performance metrics for each problem rather than revise one or more approved methods for each problem. The metrics can be applied to any candidate methodology, and provide all the evaluation criteria needed to determine whether or not it is suitable for the applicable task.

For marine pigment concentrations, the community met part of the performance-based burden, because it agreed on an accuracy metric for Chl *a* concentration, but there was no consensus for any other pigment or criteria other

than accuracy. Consequently, the SeaHARRE participants arbitrarily adopted the Chl *a* metric for all data products, and developed a set of performance criteria for all the pigments, which are presented in Table 2 as an example of what an approach based on performance metrics might look like. The four different category labels were selected for convenience, and simply provide a scale of capabilities. In some cases, this score might coincide with one of the chosen categories, like “semiquantitative,” but in other cases there might be reasons for a separate “validated” category. This language was not part of the HPLC work, because the use and application of HPLC methods is more extensive than the narrower ocean color (marine pigment) perspective adopted for the SeaHARRE activities.

Each category in a performance metric is assigned a weight and score, so the overall capability of a method is based on summing the applicable weights for each performance parameter, dividing by the number of parameters, and comparing the result to the category scores. This process permits any method to be evaluated against a) another method that is already properly validated, and b) the stated requirement for the type of work being pursued. For example, if product validation requires “semiquantitative” data, then a method with an overall score of 1.5 or more, would be suitable for the task. The classification could

also be recorded when data are submitted to a database, so future users could use only those data in keeping with their research objectives. In other words, if only “state-of-the-art” data are applicable, then the database can be searched for only this quality of data.

As long as there is some range in performance thresholds and they are set so most methods qualify for the middle portion, the use of metrics allows analysts to understand which criteria associated with their individual methods need to be enhanced to advance the overall capability of their method. In some cases, this will be a rather straightforward exercise of discovering which procedures can be improved by using more accurate components or techniques; in other cases, the development of new approaches might be needed to overcome long-standing limitations. The latter represents new research that might not occur in the absence of a performance requirement.

Because performance metrics provide a quantitative assessment of quality, they can be used to establish what constitutes a properly validated capability within each subcategory (e.g., calibration) or across an entire method (e.g., the HPL Method). Establishing the individual parameters and scores is by necessity quantitative, but the details of the underlying work remains hidden at the scoring level. Consequently, the step-by-step *best practice*, which is frequently presented in a protocol, is not available in a strictly performance-based approach. This might not be considered a significant loss of information for experienced analysts, but for new practitioners, it represents an important reason for maintaining detailed protocols. Protocols, uncertainty budgets, and performance metrics represent a natural progression: protocols are a tool to define the uncertainties (as expressed in this document), that will eventually lead to performance metrics, that will ultimately allow for any protocol to be exclusively evaluated by the metrics.

## 2.4 Provide Access to High-Quality Data

Ensuring access to high-quality data is an ongoing requirement that is expressed as a diversity of tasks in a variety of program functions. The diversity is driven by the concept of “access,” but within the context of facilitation. What this means is any group with the responsibility of providing the scientific community with access to the data needed to fulfill a set of research objectives must be prepared to facilitate the procedures used to gain access.

In a straightforward sense, access to *in situ* data is satisfied with a simple archive of all the data sets relevant to the entire calibration and validation activity (Hooker et al. 1994). Inevitably, the utility of an archive is best exploited by including the retrieval of historical data sets, which is quickly followed by the implementation of quality control, documentation, and cataloging procedures. These enhancements place continuing programming requirements

on maintaining a suitable database structure for the evolving complexity of the archive, as well as the evolving sophistication of the user-friendly interface (Werdell et al. 2003).

Access to high-quality satellite data is usually a more sophisticated undertaking, because of the volume of information involved. When more than one satellite is in operation, or if multiple missions have been archived, an efficient mechanism is needed to acquire, process, and reprocess the data. Remembering that the proper initialization of the first data are collected after launch, plus the continuing application of the atmospheric correction algorithm are inherent functions. The corresponding products must be distributed and archived, all the while being temporally monitored for any signs of satellite calibration problems. Ultimately, the user community is only satisfied if an adequate capability to browse and order data is also available.

The programming needs for facilitated data access can extend to a variety of sophisticated requirements. The SeaWiFS Project, for example, determined satellite data is best exploited if user-friendly, end-to-end processing tools for the most common computer systems were made available by the Project. The outgrowth of this undertaking was the SeaWiFS Data Analysis System (SeaDAS), which is a simplified version of the operational processing system (Fu et al. 1996 and Baith et al. 2001). SeaDAS was primarily supported by separate funding from the OBB Program, but required additional Project involvement and resources (hardware, system administration, etc.). SeaDAS would not have been possible, however, without a close working relationship between the SeaDAS team and the Project staff who had a detailed understanding of the multiple levels of processing codes, which are incorporated into SeaDAS.

One lesson from the DARR activity discussed earlier was instrument manufacturers are not necessarily the most reliable sources for high-quality *in situ* data processors. Furthermore, the experiences gained in extensive field campaigns, like the AMT program (Aiken et al. 2000), showed *in situ* data are as likely to require archival reprocessing as satellite data. In the case of the AMT data, this was driven by the evolving understanding of instrument uncertainties, like the characterization of immersion factors. One way to ensure the data in an archive are kept at the same quality level is to reprocess all the applicable data once this need is apparent. Developing *in situ* data processors, first for AOPs and then for inherent optical properties (IOPs) is not as daunting a requirement as it might first seem, because some calibration and validation data already have a single point of processing (and, thus, reprocessing). The SeaWiFS Photometer Revision for Incident Surface Measurements (SeaPRISM) data, for example, are processed and made available by the Aerosol Robotic Network (AERONET), which maintains sun photometer sites around the world (Holben et al. 1998).

Another way to ensure the data in an archive are of equal quality is to have only one data source. HPLC pigment analyses for many NASA investigators, for example, are provided by a single laboratory, as part of a contracted (but competed) service that began during the SIMBIOS time frame (Van Heukelem et al. 2002). A potential pitfall of this approach is any biases or problems associated with the central facility will corrupt a significant amount of data. An accepted way to prevent the likelihood of the latter is to make sure the chosen laboratory has a robust QA capability (including traceability to the proper standards, participation in analysis round robins, etc.) and satisfies an agreed upon performance metric†.

## 2.5 Manage Vicarious Calibration Sites

As discussed earlier, the SeaWiFS and MODIS vicarious calibration strategies were based on a single site in conjunction with accurately tracking sensor stability using onboard (lunar and solar) calibration techniques. The viability of this approach depends on additional assumptions such as knowledge of the polarization sensitivity and how it changes over time. In the case of SeaWiFS, the polarization scrambler minimizes this effect and because the fore optics (telescope, scrambler, and half-angle mirror) are protected by the telescope housing, changes in polarization sensitivity are likely to be small. In the case of MODIS, which has significant polarization sensitivity, the scan mirror is exposed and the reflectivity of the mirror has changed appreciably. This implies that the polarization sensitivity has also changed, to some degree; other properties like response versus scan (RVS) have also changed.

The ability to compare contemporaneous SeaWiFS and MODIS observations (Franz et al. 2005) allowed refinements to the polarization and RVS corrections for MODIS that would not have been possible without SeaWiFS, although no methodology for tracking changes in MODIS polarization sensitivity has been developed. For VIIRS, there will only be the onboard solar diffuser (plus stability monitor), very limited lunar calibration data, and whatever field data are available to track sensor performance. Given that effects like polarization sensitivity are a strong function of solar and sensor viewing geometries, a single calibration site will bias the vicarious calibration as a mean value for the specific range of geometries associated with the latitude of the site. Consequently, a network of sites are needed to span the full range of latitudes being observed. Because the calibration of the spaceborne sensor needs to be accurate at the 0.2% level, the calibration sites must also be intercalibrated at this level.

---

† Originally, the Center for Hydro-Optics and Remote Sensing (CHORS) provided HPLC analyses for the OBB Program, but during SeaHARRE-3, significant problems were discovered with the CHORS method. Since then, HPL has been contracted for these services.

For sensors that do not tilt, MODIS showed the location of a single calibration site at a low latitude, even for non-noon orbits, is severely affected by sun glint during the summer months, which reduces the number of vicarious calibration matchups. While sun glint can be modeled and removed rather well for moderate contamination levels, it is highly polarized, making accurate sensor characterization even more essential. In the present SeaWiFS and MODIS processing, much of the glint-contaminated portions of the data are retrieved, but MODIS saturates in the NIR bands making the data unusable. It should be noted that the VIIRS 748 nm atmospheric correction band has a single gain and will saturate in sun glint.

One approach to achieving the needed level of consistency in vicarious calibration is to have a field network that uses a common instrument design, a common deployment strategy, a common data analysis methodology, and common calibration standards. The latter should include the active involvement of qualified personnel from the institution maintaining the calibration standards, and the network should be managed by the same group to ensure these practices are enforced. The AERONET is a good example of a large network maintained under the stewardship of one group but with an international participation. The distribution of sites requires local logistical support, access to sites, routine site and instrument servicing, etc. These collaborations need to be resolved at the international agency level with formal agreements. An international steering or advisory group, which includes representatives from the nations hosting the sites and their contributing space agencies, is also needed.

Regardless of the chosen vicarious calibration approach, all of the collected data need to be publicly available, for example, through the SeaWiFS Bio-Optical Archive and Storage System (SeaBASS). This transparency is needed not only to permit other investigators to scrutinize the quality of the data, but also so they can pursue alternative research inquiries. The development of new algorithms and their attendant data products, for example, might require a different vicarious calibration procedure, perhaps tied to a particular atmospheric correction scheme and radiation transfer code, than what the overall activity is using. In general, it is always rewarding to facilitate the use of diverse techniques, because intercomparison of all applicable methods usually leads to the discovery of problems whose correction improves the overall capability.

## 2.6 Address Optical Complexity

A significant challenge for future ocean color research will be to maintain the level of success achieved in deep-ocean (Case-1†) waters in the coastal ocean and marginal

---

‡ By definition, the optical properties of Case-1 waters are solely determined by the phytoplankton and its derivative products (Morel and Prieur 1977), whereas Case-2 optical properties are also determined by other material, e.g., from terrestrial or bottom origin.

seas, which means the influence of dissolved and particulate constituents will be increasingly important. In these more predominantly Case-2 waters with significant vertical structure and optical complexity, AOP above-water methods are likely to be superior to in-water AOP measurements, because they can measure the surface layer directly without instrument shading, and surface perturbation problems are easily removed—although platform shading and reflection sources must be properly minimized (Hooker and Zibordi 2005).

Dissolved and particulate constituents were not a significant part of the Case-1 investigations discussed above, because SeaWiFS and MODIS were designed for Case-1 waters, and the principal water components were assumed to covary; in fact, in keeping with the definition of Case-1 waters, the primary contribution was simply the marine phytoplankton. This does not mean the large effort expended in the open ocean is not applicable to the follow-on problem set. Indeed, the proper perspective is to use the lessons learned in the optically simpler Case-1 environment to establish a starting point, or foundation, for investigating the optically more complex Case-2 environment.

## 2.7 Develop and Evaluate Instrumentation

As mentioned earlier, the quantification of uncertainties demands an evolving understanding of measurement uncertainties, which requires detailed evaluations of instrument characterizations. This evolution leads naturally to the development of modified designs from existing technologies or completely new concepts, which steadily advance the state of the art. The former is potentially the more attractive, because if commercial off-the-shelf hardware (COTS) can be successfully adapted to a calibration and validation requirement, a significant amount of development cost (and time) is saved.

The SeaWiFS and SIMBIOS projects invested directly in instrument evaluation and development, which led to many important accomplishments. Two of the more notable, which are directly applicable to the plan presented here, are the aforementioned SeaPRISM capability and the *Bouée pour l’acquisition de Séries Optiques à Long Terme*† (BOUSSOLE) project. Both are examples of using COTS hardware for calibration and validation requirements.

SeaPRISM is a modified fully-autonomous, commercial sun photometer, used by AERONET, which measures the sea surface and sky after performing the normal sun and sky measurements needed for sun photometry. A prototype unit was developed in partnership with the Joint Research Centre‡ (JRC) and field commissioned at the *Ac-*

*qua Alta* Oceanographic Tower (AAOT) in the northern Adriatic Sea (Hooker et al. 2000b). The prototype was assessed for the validation of remote sensing radiometric products in coastal waters (Zibordi et al. 2002b), and a one-year time series of data were compared with simultaneous in-water measurements for a wide variety of sun elevations and environmental conditions. The average relative percent difference§ (RPD) values between the above- and in-water  $L_W(\lambda)$  determinations were less than 2% in the 412–555 nm spectral interval (Zibordi et al. 2004).

The good validation results achieved with SeaPRISM led to a separate investigation of using the instrument for vicarious calibration. The gain factors computed from a one-year demonstration phase data are presented in Table 3 (remembering that the gain factors are the aforementioned adjustments to the responses of the SeaWiFS channels to force agreement with the *in situ* data). The SeaPRISM prototype did not have a complete overlap with the satellite channels, because AERONET required a certain minimum number of sun photometer wavelengths (this restriction has been dropped). The close agreement between the SeaWiFS vicarious gains computed from SeaPRISM and MOBY data (even though the SeaPRISM data are from a coastal deployment site), shows how well a low-cost alternative methodology based on COTS hardware *might* work, but it is not the only applicable example.

**Table 3.** The vicarious calibration gain factors for SeaWiFS determined using SeaPRISM and MOBY. The relative percent difference, RPD, is computed with MOBY as the reference value.

Wavelength [nm]	Gain Factor		RPD [%]
	SeaPRISM	MOBY	
412	1.0462	1.0360	0.98
443	1.0129	1.0126	0.03
555	0.9999	0.9939	0.60
670	0.9816	0.9627	1.96

Unlike SeaPRISM, the BOUSSOLE project was specifically established to collect vicarious calibration data and is an alternative in-water buoy developed in partnership with the *Laboratoire d’Océanographie de Villefranche*¶ (LOV). The innovative aspects of the design include a taught-cable mooring, which does not require a separate surface flotation buoy, and a tubular *transparent-to-swell* superstructure ensures a minimal shading perturbation from the mooring plus a very stable mounting system for the instruments (Antoine et al. 2006).

† Literally translated from French as the “buoy for the acquisition of a long-term optical series.” “*Boussole*” is the French word for “compass.”

‡ The principal investigator (PI) who worked with CIMEL Electronique (Paris, France), the manufacturer of the sun photometer, to make sure the sampling protocol would produce the highest-quality data was Giuseppe Zibordi.

§ The RPD is defined as  $100(X - Y)/Y$ , where  $X$  is an independent observation,  $Y$  is the dependent reference value, and the factor of 100 yields units of percent.

¶ The PI who conceived the buoy and worked with Satlantic, Inc. (Halifax, Canada), the manufacturer of the radiometers, to make sure the optical measurements would be of the highest quality was David Antoine.

Figure 3 presents matchups from the BOUSSOLE optical data with three ocean color satellites: the European Medium Resolution Imaging Spectrometer (MERIS), SeaWiFS, and MODIS-A. MERIS has never been vicariously calibrated, so these data show some biases, particularly in the red domain. The other two satellite sensors were vicariously calibrated using MOBY data, and the BOUSSOLE results exhibit almost no bias, although a small amount is seen in the red wavelengths. This is a significant result, because BOUSSOLE is completely independent of the MOBY activity, except for the radiometric traceability to the NIST standard for spectral irradiance. The BOUSSOLE calibration gains are in good agreement with the MOBY gains: to within 0.6% for 443–670 nm and to within 3.4% for 412 nm. The poorer agreement at 412 nm might be caused, in part, by undersampling at this wavelength, because only one set of in-water radiometers are equipped with this channel.

The BOUSSOLE accomplishment is placed in an even more remarkable context when the level of agreement is considered with respect to the coefficient of variation in the MOBY gains for SeaWiFS (based on two standard deviations), which is approximately 0.5% (Franz et al. 2007). This means almost all of the BOUSSOLE data are at a quality level that is largely indistinguishable from MOBY (412 nm being a notable exception) even though COTS instrumentation was used. The potential of humble approaches like BOUSSOLE and SeaPRISM suggest straightforward and low-cost instruments might be viable alternatives for vicarious calibration measurements as long as their capabilities properly scale to the performance and characterization of the satellite sensor.

### 3. The Proposed Activity

The emphasis here is on producing an interconnected plan wherein each individual piece of the activity is represented, and its relationship to all the other parts of the biogeochemistry program is made clear. Based on the experiences derived from prior satellite missions, a horizontal organizational scheme is imagined. The overall leadership comes from the OBB Program Manager, and the entire enterprise is split into two components with equal stature: calibration and validation, plus satellite data processing. A pictorial representation of this concept is shown in Fig. 4, which is based primarily on a macroscopic view emphasizing how the elements are organized, but additional finer detail is also shown. The P indicators are a reminder that the denoted element requires an agreed upon and published set of sampling, analysis, and data reporting protocols. The protocols must include performance metrics with accuracy, precision, and QA thresholds that establish the criteria for a) routine research, b) product validation, c) product refinement, and d) satellite calibration (if applicable).

The details of Fig. 4 are based on the basic tasks of the two main components plus the current objectives of

the Carbon Cycle and Ecosystems Roadmap†. The former is distinguished by an *internal* core set of responsibilities and the latter is introduced through an outer *external* connecting-core ring of competed and contracted activities. The dates shown with some components indicate when they are expected to join the whole enterprise.

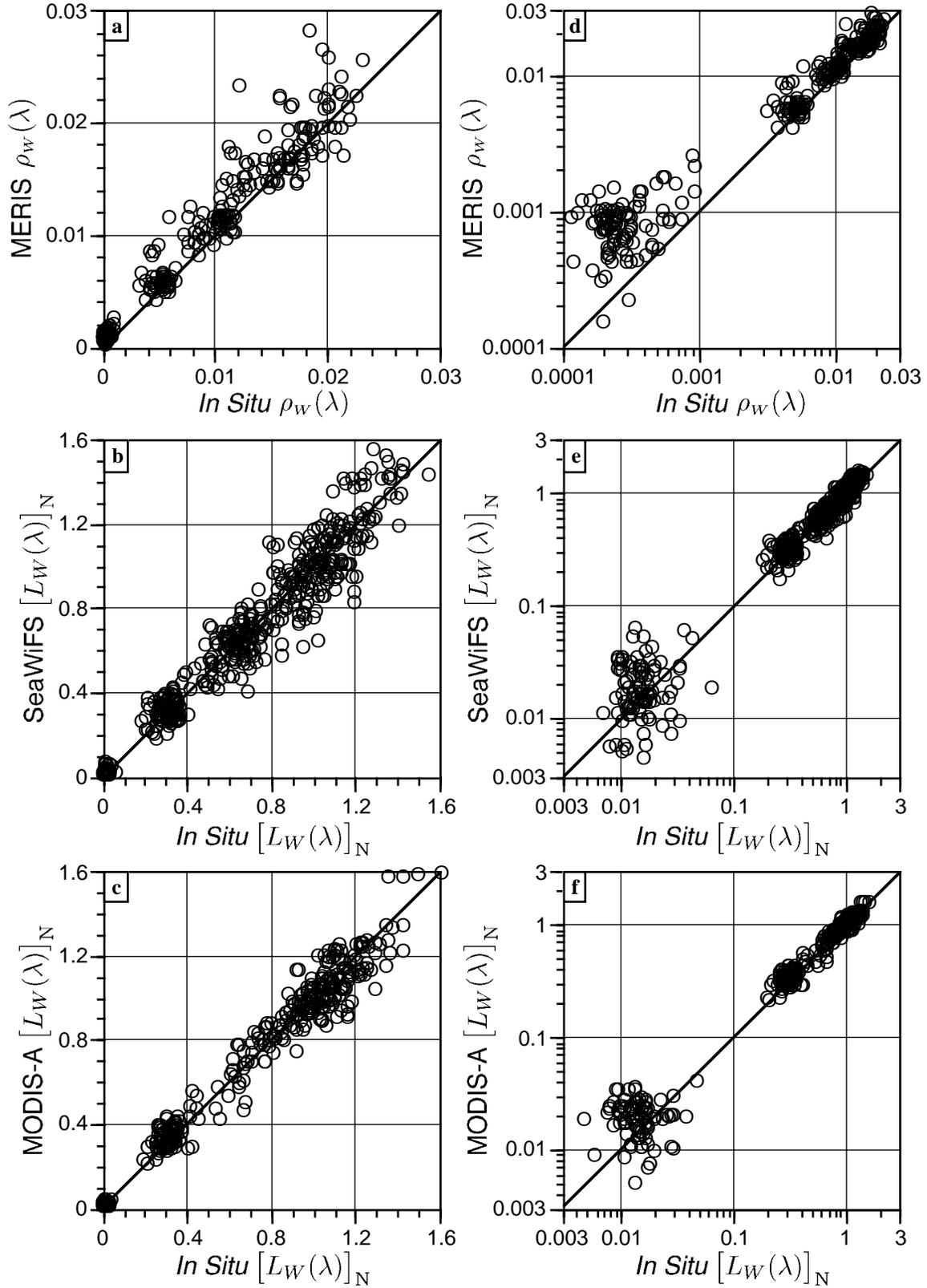
Depicting the various parts as interlocking, equally-sized pieces emphasizes the interdependence of all the elements, plus the fact that no one part is assumed more important than any other. The interdependence is not just associated with one task relying on another for successful execution. What is imagined here is that the expertise resident within the internal set of core functions will extend outward to the connecting-core ring to ensure every element has an alternative execution capability if the prime capability fails (for whatever reason). Note that this philosophy provides a gradation of effort in each element, because it can be fully funded (with full representation by an externally competed representative) or partially funded (with part-time representation by a core functionary).

The day-to-day management of the activity will be the responsibility of a Calibration and Validation Chair and a Satellite Data Processing Chair who will oversee their respective components. These two chairs will be part of a Calibration and Validation Team that will provide expert advice to the OBB Program Manager who will be in charge of the entire activity. The other team members will be selected from the connecting-core competencies already discussed: standards and traceability, biogeochemistry, AOPs, IOPs, *in situ* database, vicarious calibration, product validation, and atmospheric correction. New positions will be created as new science topics are added (e.g., carbon abundance, primary productivity, etc.) and positions will be deleted as specific elements are completed or suspended.

The evolution of funding levels and programmatic priorities will necessarily alter the idealized implementation and temporal realization of any plan. The delineation of core and connecting-core elements provides a program manager with a unique opportunity to understand which parts are the essential elements of the activity while tactically implementing individual pieces (or portions thereof), which expand or contract the overall scope of the activity as differing budget and funding opportunities materialize. This should not be interpreted as advocating a predetermined set of budget-minded principles—like protecting the core functions at the expense of external activities; the

---

† The Carbon Cycle and Ecosystems Roadmap is focused on the implications of environmental change and human activities on the Earth’s ecosystems and the biogeochemical cycles that are critical to the habitability of the planet in terms of food production, sustainable resource management, carbon management, conservation of biodiversity, and the maintenance of a healthy environment. A discussion of the science questions associated with these topics is available at the <http://oceancolor.gsfc.nasa.gov/DOCS> Web site.



**Fig. 3.** Satellite matchups at the BOUSSOLE site for **a)** MERIS  $\rho_w(\lambda)$  values (412, 443, 490, 510, 560, 670, and 683 nm); **b)** SeaWiFS  $[L_W(\lambda)]_N$  values (412, 443, 490, 510, 555, and 670 nm); and **c)** MODIS-A  $[L_W(\lambda)]_N$  values (412, 443, 488, 510, 551, 670, and 683 nm). The solid line is the 1:1 line. Logarithmic scales for panels **a-c** are shown in panels **d-f**, respectively, in order to magnify the low values in the red domain.



point here is to simply admit that the annual capabilities of the plan are a function of the funding available and any rescope (led by the OBB Program Manager with advice from the Calibration and Validation Team) should be done with an understanding of the short- and long-term interdependencies of the various elements.

### 3.1 Satellite Data Processing

The execution of the SeaWiFS and MODIS missions demonstrated one of the most important aspects of the calibration and validation process is a strong link with the data processing system. This is only possible if both components are properly organized, staffed, and integrated with one another at the core level of responsibilities. For the satellite data processing component, the latter includes the following tasks:

- Acquiring, distributing, and archiving data products (including continuity products associated with climate analyses);
- Processing the satellite data from the lowest to highest levels<sup>†</sup> from all the applicable satellites (which might include historical missions);
- Participating in the temporal monitoring of the satellite sensor’s performance; and
- Reprocessing all the data products when changes in the scientific or engineering understanding of the data dictates it.

Note that the last item occurs, and is scheduled, only as a result of a continuing dialogue with the scientific community, and is not done in isolation.

The design and construction of a data processing system capable of satisfying the core tasks given above is not the subject of this document, but it is worth noting what architectural functions are required, because they are immediately applicable to the final set of core and connecting-core tasks. Using the SeaWiFS processing system as an example (McClain et al. 2004), the five key functions are: a) processing large volumes of data in a timely fashion, b) automating as many of the tasks as possible, c) allowing changes to the processing methods to be easily implemented, d) accommodating multiple processing streams, and e) providing easily understood documentation and interfaces. Although the successful implementation of these functions do not always directly relate to the connecting-core tasks, they provide the needed degree of efficiency to

permit the people involved to spread their expertise outward through the connecting-core elements.

Again using the SeaWiFS experience as a model, the connecting-core elements that the core scientists have participated in are:

- Applying the atmospheric correction algorithm;
- Vicariously calibrating the satellite sensor;
- Validating the bio-optical algorithms and the associated data products;
- Incorporating new sources of applicable data (e.g., *in situ* SeaPRISM observations from AERONET and remote sensing measurements from CALIPSO);
- Maintaining the *in situ* database (SeaBASS) and providing satellite overpass schedules and near-real time data support to field campaigns;
- Implementing new global and regional data products; and
- Refining the image processing software (SeaDAS).

This list represents a diverse set of skill and knowledge sets and firmly establishes the viability of relying on a core set of capabilities that connects all the programmatic pieces together. Just as importantly, it also shows the wider community can be properly joined to the basic functions through a connecting-core interface.

The reason a strong link with the data processing system is so critical is the amount of analysis that must take place to maintain the accuracy of the data products. While simple evaluations based on a few scenes can be easily conducted on small systems, analyses spanning an entire mission on a global scale can only be conducted efficiently and quickly on a *main* processing system. Batch processing on this scale for calibration and validation purposes is possible because a properly designed processing system—like that built for SeaWiFS and, ultimately, MODIS—has all the level-0 satellite and ancillary data online and can be easily configured for customized analyses.

The processing system designed from the earliest stages of the SeaWiFS Project accommodated offline processing in support of the calibration and validation element (McClain et al. 2004). This flexibility permitted the incorporation of MODIS sea surface temperature processing as an independent (parallel) processing stream without any interruption in the routine ocean color processing. Prior to each SeaWiFS (versions 1.0–5.1) and MODIS (versions 1.0–1.1) reprocessings, extensive testing of algorithm refinements were conducted (e.g., revised MODIS polarization tables, updated ozone ancillary data, and inclusion of the bidirectional correction). At present, tests are normally run on both SeaWiFS and MODIS with the same algorithms, so the results can be compared across both missions. The capability is facilitated by using a standard processing code, *MSL12*, across all sensors. These tests provide complete information on global and regional biases

<sup>†</sup> The extent of data processing—usually referred to as the *level*—begins with the (raw) data received on the ground, which is denoted level-0. Calibrated and geolocated data (level-1) are used to derive the geophysical (level-2) products (e.g., the chlorophyll *a* concentration), which are usually space-time averaged and binned onto a standard grid (level-3). The level-1 and level-2 products include metadata identifying the ancillary fields required for level-2 processing. Complete details for each level are available in the *SeaWiFS Technical Report Series* (Firestone and Hooker 2004).

and trends that cannot be identified based on more limited analyses. Within the first two-years of the SeaWiFS processing group having the responsibility for MODIS processing, over 80 such time-series analyses were executed, which led to major improvements in both data sets.

The procedure for setting up the tests is straightforward. The calibration and validation lead analyst provides the modified code, plus the input parameters and tables in a prescribed format, to the processing system lead. Typically, the time-series tests use four consecutive days of each month. The processing system is automated using databases to control the processing sequence and data handling. Once the processing is initiated, it takes only a few hours to complete, including the level-3 processing. Once the level-3 products are available, various standard global and regional time series and statistical analyses are computed and posted for the calibration and validation group to inspect. This level of cooperation and responsiveness is best achieved by having the calibration and validation group collocated and under the same management structure as the data processing group.

Throughout the time period of the SeaWiFS mission, reprocessings have been executed approximately once every 18 months. In each case, the calibration and validation group posted the results of the tests along with the recommended algorithm revisions for community comment prior to the actual reprocessing. The generation of science quality data that incorporates the new understanding of sensor calibration, atmospheric corrections, and bio-optical algorithms necessitates these reanalyses. The field of ocean color remote sensing is expanding rapidly as a result of the SeaWiFS and MODIS missions, so the OBB Program must provide state-of-the-art data sets that incorporate these advances in a timely manner. The only way to achieve this is to have a data system that is designed to be adaptable and responsive to the calibration and validation requirements. An additional benefit is that the SeaDAS system is managed as part of the data processing system and incorporates all the refinements of each reprocessing in synchrony with the reprocessing (i.e., updates of SeaDAS are released with each reprocessing to ensure consistency).

### 3.2 Calibration and Validation

The calibration and validation of an ocean color satellite includes spacecraft, atmospheric, sea surface, subsurface (or *in situ*), laboratory, and data analysis tasks, all of which require pre- and postlaunch activities. The most important goal of this effort is to produce water-leaving radiances within an agreed upon uncertainty level and chlorophyll *a* concentration range (hereafter referred to as the *satellite performance metric*). The partitioning of the necessary work plan into its functional parts can be accomplished several different ways. The approach adopted here is based on the experience acquired during the execution of the SeaWiFS mission and follows directly from Fig. 4.

For recent remote sensing missions, the satellite performance metric requires field instruments with a calibration and measurement capability in keeping with a 1% radiometry requirement, so the quadrature sum of uncertainties is to within the overall uncertainty budget. Because these challenging *in situ* measurements will frequently be acquired from a variety of field instruments over the mission lifetime, a measurement assurance program is required. This program consists of several activities:

- An accurate prelaunch characterization and calibration of the spaceborne instrument;
- Multiple vicarious calibration sites in clear waters and atmospheres to provide time series of water-leaving radiances for postlaunch vicarious calibration (which hopefully includes international partnerships to expand the global coverage, especially into the Southern Hemisphere);
- Clearly defined protocols with performance metrics for established data collection methodologies;
- Direct comparison to the appropriate national standards laboratory (NIST in most cases) to verify uncertainty budgets;
- Managing the development and evaluation of instrumentation to ensure measurement uncertainties are quantified and minimized; and
- Temporal monitoring, quality assurance, and data analysis procedures for tracking the postlaunch performance of the satellite sensor and the validity of the derived products.

The net culmination of many of these activities is the deployment of the instruments and methodologies on sponsored and interdisciplinary field campaigns. In this context, “sponsored” refers to joint-agency expeditions (e.g., NACP or SOCP) and “interdisciplinary” stresses the new paradigm for calibration and validation wherein optical, biological, and chemical expertise are deployed with equal stature (the original paradigm emphasized primarily optical measurements). Under this plan, the Calibration and Validation Activity will help coordinate the needed field campaigns, publicize and negotiate ship time opportunities as well as atmospheric and aerosol characterization experiments, and facilitate the sampling needs of new science topics.

Although sponsored cruises will be part of research announcements (usually with broad science objectives), separate campaigns with focused calibration and validation objectives—but with competed participation—will be needed. These specific experiments will be organized to collect the optical, biogeochemical, and atmospheric data required to address the underlying question or hypothesis while properly characterizing the environmental conditions of the experimental site. The point here is to establish the tenets of the inquiry, choose the experimental location, provide the sampling platform (a ship, offshore structure,

airplane, etc.), and then select the participating scientists from a competed announcement.

It is natural to think of many field campaigns in terms of ship transects, but time series will be one of the most important data collection opportunities. The advantage of a time series is it provides a proven pathway for quantifying the quality of the data being archived, but only if it extends over a significant amount of time (in terms of the time scale of the underlying processes). Indeed, for climate-quality data, time spans on the order of 10 years or more are rather easily imagined. Lengthy field work is usually very expensive, so community support and accountability will be essential for such activities (Sect. 4.1).

Coastal ocean experiments will be designed to evaluate and refine instruments and protocols used to sample and analyze optical and biogeochemical measurements, particularly those involving issues unique to the coastal zone (plumes, shallow water, vertical complexity, turbidity, etc.). Major field campaigns will be organized or synchronized with other institutional partners and government agencies (when possible) as new programmatic elements are implemented by the OBB Program (e.g., vicarious calibration sites and the SOCP campaign). The management of the vicarious calibration sites will involve coordinating ship time for routine maintenance (instrument calibration and exchange, etc.) as well as ensuring that IOP and biogeochemical validation data are collected and processed according to the established protocols.

Calibration and validation activities will be aligned to meet the requirements of new science topics as they are introduced: oceanic carbon abundance, coastal carbon abundance, oceanic primary productivity, and physiology and functional types. Coordination of the field campaigns and protocol evaluation experiments associated with the new science topics will be pursued before the topic is introduced to ensure a mature capability is available when it is needed most. Protocol issues that need to be resolved include a) choosing a method† and protocol that is the most appropriate for quantifying gross or net primary productivity, and b) developing a carbon reference material for use when measuring particulate organic carbon (POC). A possible solution for the latter might be modeled after what is being done to verify measurements of dissolved organic carbon (DOC). Over the past several years, the scientific community has relied on a consensus reference material (deep water from the Sargasso Sea) distributed by the University of Miami (Department of Marine and Atmospheric Chemistry). Protocols for current and new optical instrumentation will be required to provide a rigorous measure of quality and consistency for the *in situ* database (e.g., scattering sensors, fluorometers, etc.).

---

† Candidate methods include  $^{14}\text{C}$  or  $^{13}\text{C}$  incubations, enriched or natural abundance oxygen isotopes, plus classic light- and dark-bottle incubations.

As the ocean color community continues development of algorithms for IOP satellite products—particle and colored dissolved organic matter (CDOM) absorption or particle backscatter, for example—the quality of the validation data becomes more critical. This is particularly important when the IOP data are used to derive biogeochemical products (e.g., deriving the POC from beam attenuation). Protocols generated from workshops sponsored by NASA in the late 1990s for particle and CDOM absorption (Mitchell et al. 2000 and 2003) should be updated to promote consistency in protocols and data processing, for analysis of samples collected in coastal waters, and to take into account advances in instrumentation (e.g., liquid capillary waveguide absorption instruments). Commercial vendors continue to develop new in-water IOP instruments (absorption, attenuation and scattering sensors, as well as fluorometers). Measurements from the new instruments should be compared with the previous instruments and with discrete bench-top measurements. This will require IOP instrument and analysis round robins.

### 3.3 Competed Elements

An integral part of the plan is to use the peer-reviewed process to fill the majority of the connecting-core elements; a small number of these elements will be filled using contracts where appropriate. In keeping with the duality of responsibility in connecting-core elements (internal and external scientific participation), if budget levels do not permit external participation for all of these positions, some of them can be filled by internal (core) scientists (if deemed appropriate by the OBB Program Manager) to ensure a minimum level of representation.

The competed (and contracted) activities encompass a broad range of scientific topics and tasks: a) atmospheric and aerosol characterizations, including improvements and new approaches for atmospheric correction algorithms; b) improved or new data products (global and regional algorithms for AOPs, IOPs, biogeochemical constituents and processes, etc.); c) characterization of new satellite sensors (VIIRS and future sensors); d) sponsored and interdisciplinary field campaigns (NACP, SOCP, vicarious calibration sites, etc.); and e) fundamental scientific research consistent with the Carbon Cycle and Ecosystems Roadmap and relevant to calibration and validation. Additional areas of endeavor in keeping with NASA programmatic objectives will be added when necessary.

Improving atmospheric correction algorithms will require advances in modeling, as well as field activities to better characterize the atmosphere and aerosols. Current efforts include utilizing the 1.24 and 1.64  $\mu\text{m}$  bands on MODIS and incorporating new data streams from satellite sensors (e.g., CALIPSO) and surface instruments (SeaPRISM). Development of regional algorithms for current or new ocean color data products will be needed to meet

NASA objectives in coastal and deep-ocean waters. Advances in the next generation of ocean color satellite sensors, such as hyperspectral capability and bands in the near ultraviolet (UV), will necessitate calibration and validation field campaigns, as well as the development of new algorithms and models. Contracts with certain laboratories to conduct biogeochemical analyses (e.g., HPLC pigments, POC, DOC, etc.) may be used to maintain a consistent uncertainty level for the *in situ* measurements (as is being done presently for HPLC pigments).

A recolored version of Fig. 4, wherein the competed (green) and shared (yellow) connecting-core elements are explicitly denoted, is presented in Fig. 5. The latter includes competed, contracted, and internal contributions, the mixture of which is determined largely by the topic type. For example, the calibration and validation component is expected to be involved in the management of the vicarious calibration site(s), which means it will also be involved in the optical properties element (at least the AOP portion), as well as the standards and traceability element. Almost all of the satellite data processing connecting-core elements are shared responsibility elements, because there has to be an internal competence to interact with the community contributions in these elements.

Calibration and Validation Team members will a) establish and review calibration and validation requirements for the vicarious calibration sites, field campaigns, and new science topics; b) plan field experiments to address pertinent calibration and validation issues (e.g., AOP, IOP, and biogeochemical data collection in Case-2 waters); c) maintain up-to-date protocols for their respective elements; d) assist with coordinating instrument and analysis round robins; e) review compliance of core, competed, and contracted activities with established Protocols; f) enforce accountability and transparency for the wider scientific community; and g) provide recommendations to the NASA program manager regarding calibration and validation activities, including field campaigns and new science topics. The Team will meet (physically or via conference call) quarterly or as needed to satisfy its obligations.

## 4. Issues and Discussion

The development of the plan presented here focused on the lessons learned from ocean color missions already successfully planned and launched. At first glance, this might appear to be a rather myopic perspective, but the future for ocean color—in terms of satellite missions—is currently very similar to the past. The satellites approved for launch are all fixed-wavelength sensors with very similar spectral, orbital, and performance specifications to what has already been flown. At some point, however, this will not be true, and the types of scientific investigations and methods needed to satisfy the Advanced Science Plan will very likely require different satellite designs, which, in turn, will require alternative field studies and instrumentation.

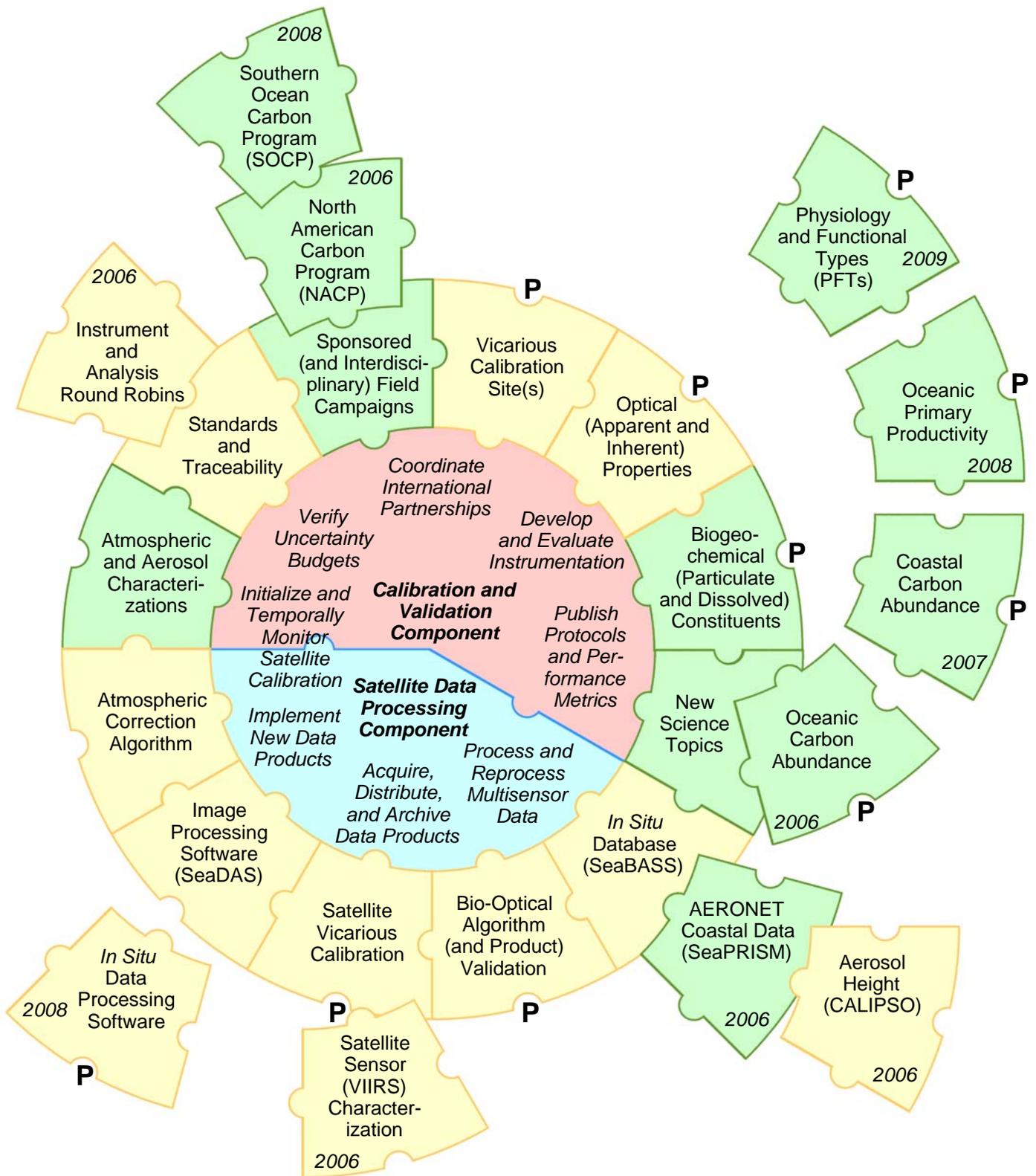
Properly forecasting what is needed for a long-term plan automatically requires choices between competing visions. Although the authors of this plan have done their best to make sensible choices—at least in terms of the here and now, plus the lessons learned from prior missions—*a priori* there is no way to know if these choices will withstand the test of time without initiating the entire process. Nonetheless, some safeguards can be put into place, which is to say an accountability procedure can be defined, and some of the most likely issues associated with the perspective adopted here can be discussed.

### 4.1 Accountability

One aspect of past missions that changed from launch to launch was how the scientists involved were held accountable for successfully completing their various tasks. The plan described here will be one of the substantial aspects of the OBB Program, so accountability is a required component. The plan is based on uncompleted (core) as well as competed and contracted (connecting-core) tasks, and the procedures for fulfilling the latter provide a certain amount of built-in accountability (i.e., peer review), but implementing the former basically does not. The lessons learned from prior missions—especially MODIS—establish the necessity of an independent core group to provide objective performance evaluations of calibration and validation activities. One of the best mechanisms for maintaining objectivity is to have annual reviews of the science and deliverables involved.

The first level of accountability resides with the OBB Program Manager who will be in charge of the overall activity. A second level of accountability is provided by two upper-level oversight opportunities: a) the Calibration and Validation Team, and b) an annual external review of the entire enterprise. The annual review will consist of teleconference reviews of the (contracted) connecting-core elements and a public review of the core elements with the broader ocean color community invited. The two types of reviews are envisioned to save on travel and logistics while maintaining an effective review capability. If there are important issues or results within the connecting-core tasks that need public input, separate presentations will be scheduled during the next review opportunity.

The annual review of connecting-core tasks will be overseen by the Calibration and Validation Team, whereas the core tasks will be reviewed by an Evaluation Board with participation from the public audience. The Evaluation Board Chair will be the OBB Program Manager who will select whatever representation is deemed appropriate (the composition of the board will most likely change over time in concert with the evolution in the types of science topics and technical challenges). Recommended revisions from the annual reviews will be evaluated by the OBB Program Manager, and any corrective measures will be implemented by the appropriate Calibration and Validation Chair or Satellite Data Processing Chair.



**Fig. 5.** The Fig. 4 organizational framework recolored to emphasize competed (green) and shared (yellow) elements. The latter includes competed, contracted, and internal contributions. The core elements are shown in the original blue and red color scheme.

To ensure transparency and effective communications with the community, core and connecting-core scientists will provide annual reports of their activities to the Calibration and Validation Team, which will be published as a NASA Technical Memorandum. In addition, review presentations will be made available through an appropriate Web site, and summaries of the comments and recommendations from the Evaluation Board, as well as the agreed upon corrective measures (if any), will also be posted.

A third level of accountability will be applied at the lower level of the individual elements and will involve two strategems: a) for elements with shared responsibilities between internal core functions and external connecting-core activities, the scientists involved will review and critique the work from their companion representatives; and b) the tasks associated with each element will be executed as part of one or more science objectives, which have been approved by the wider community. The point of the latter is to make sure all elements have a scientific context. It will be the responsibility of the principal scientist in each element to establish the scientific goal(s) of the element, and to present the scientific accomplishments of the element at the annual review. For data collection activities, in particular, the science objectives are seen as an essential QA opportunity to ensure all of the data are of the highest quality possible. This means *the scientists must have scientific goals that require a continuing assessment of data quality (and not a reliance on a centralized data center to detect data quality or reporting issues)*.

Lastly, there is a community-wide responsibility in this approach, as well, because the entire enterprise is only viable if the community participates and supports its evolution. In particular, the community is responsible for being directly active in many elements (to name a few):

- Participating as a member of the Calibration and Validation Team;
- Refining and approving the Protocols and performance metrics;
- Defining new algorithms and confirming the validation procedures to be used;
- Collecting field data for sponsored (and interdisciplinary) field campaigns, atmospheric and aerosol characterizations, as well as vicarious calibration sites, and submitting all these data to SeaBASS for public use;
- Enhancing the capabilities of SeaDAS to keep pace with new requirements and science topics;
- Providing expertise in the proper measurement of AOPs, IOPs, and biogeochemical (particulate and dissolved) constituents and processes;
- Improving the capabilities of the atmospheric correction algorithm;
- Taking part in instrument and data analysis round robins; and

- Ensuring the widest—domestic and international—participation of ocean color scientists.

The last item deserves additional consideration, because research is a global enterprise, and a significant part of the problem set and solutions are directly attributable to international scientists, institutes, and programs (e.g., AERONET, SeaPRISM, and BOUSSOLE are significant international accomplishments). A noted accomplishment of the SIMBIOS project was its success in uniting the international community and gaining their support in pursuing ocean color objectives.

## 4.2 Above- versus In-Water Radiometry

Although the capabilities of the above-water approach for determining water-leaving radiances have been shown to be equivalently capable as traditional in-water methods (Hooker et al. 2004 and Zibordi et al. 2004), disagreement as to the applicability of above-water methods to calibration and validation exercises persists<sup>†</sup>. At one level, this is an odd situation, because the spaceborne instrument is an above-water sensor—*so above-water radiometry is clearly a tractable problem*—but on another level, it is rather anticipated, because it takes time for scientific cultures to absorb new methodologies and change the way the scientific process is executed.

The encouraging above-water vicarious calibration results presented for the SeaPRISM prototype (Sect. 2.7) are simply an affirmation that this methodology is available for exploitation. In addition to providing an off-the-shelf alternative to an in-water technique, it is important to remember the above-water approach a) does not suffer any significant bio-fouling problems, b) does not require any self-shading corrections (although platform perturbations must be properly minimized), and c) does not have anywhere near the vulnerability of an in-water mooring (from severe weather, pleasure craft, or commercial fisherman).

## 4.3 Hyperspectral Radiometry

There are no hyperspectral satellites on orbit or waiting to be launched (VIIRS, for example, is a fixed wavelength sensor). Nonetheless, at least two potential ocean color mission concept instruments are being designed with some hyperspectral and near-ultraviolet band capability. A future ocean color satellite with hyperspectral capability may improve the characterization of oceanic constituents, but in the interim, the good vicarious calibration results achieved with SeaPRISM and BOUSSOLE—which are both based on commercially available, fixed-wavelength radiometers—suggest low-cost alternatives to vicarious calibration are

---

<sup>†</sup> In terms of the data submitted to SeaBASS, for example, the majority of the radiometric data are from in-water profilers, although the percentage contribution from above-water instrument systems is steadily increasing (currently about 45%).

accessible. Potential applications for hyperspectral sensors include improving Chl *a* algorithms in certain regions (Chang et al. 2004), resolving bottom effects on water-leaving radiances (Chang et al. 2004), detecting harmful algal blooms and accessory pigments that could be used to determine phytoplankton taxonomy (Lee and Carder 2004), and discriminating the multiple constituents that contribute to absorption and scattering properties of the ocean.

The present designs for both the Coastal Ocean Carbon Observations and Applications (COCOA) and Ocean Carbon, Ecosystem, and Near-Shore (OCEaNS) mission concepts incorporate bands in the UV and hyperspectral capability. Regardless of what actually materializes from these plans, it seems appropriate to begin the process of agreeing on revisions to the field sampling and data analysis protocols in anticipation of the initial work that will be needed to understand how best to support a hyperspectral mission. In addition, there is the need to agree on what constitutes “hyperspectral” sampling, because for many problems, a great deal of progress can be made by simply adding greater spectral diversity (i.e., many more channels) to existing radiometer designs (for example, 13- or 19-channel radiometers can be built in the same form factor used to produce 7-channel instruments).

#### 4.4 Ultraviolet Wavelengths

Improvements in atmospheric correction algorithms are essential for obtaining accurate water-leaving radiances. For clear waters, the ocean is optically black in the NIR, so these bands (765 and 865 nm for SeaWiFS) are used to estimate aerosol radiance levels and to select the most appropriate aerosol optical model in order to extrapolate the NIR aerosol correction to the visible bands. This *black-pixel* assumption for the NIR bands is not valid for productive or turbid coastal waters because of backscattering by particles. Current atmospheric correction models also tend to overcorrect for aerosols (Siegel et al. 2000, Bailey et al. 2003, and McClain et al. 2004).

In many coastal waters, absorbing aerosols and high in-water particle loads invalidate the black-ocean assumption for the NIR bands and complicate standard atmospheric correction algorithms. Negative  $L_W(\lambda)$  values for the 412 nm band (and at times for the 443 nm band) on SeaWiFS and MODIS occur frequently in coastal waters, especially for nearshore waters and estuaries. Atmospheric correction difficulties caused by absorbing aerosols extend far beyond coastal waters. For example, Saharan dust particles and biomass burning aerosols are transported across the Atlantic Ocean from Africa to the Caribbean.

The addition of UV bands can be used to flag and improve atmospheric correction algorithms in the presence of absorbing aerosols. The combination of UV and longer NIR bands (greater than 1  $\mu\text{m}$ ) may significantly improve the current black-pixel assumption for Case-2 waters. Current Chl *a* algorithms do not perform well in certain coastal

waters, because of atmospheric correction issues and the high concentrations of in-water constituents (detritus, suspended minerals, phytoplankton, and CDOM), which have overlapping absorbance spectra at blue wavelengths.

Wavelengths in the UV part of the spectrum might be exploited to distinguish the absorption signals of CDOM, Chl *a*, detritus, and minerals (due to high UV absorbance by CDOM relative to the other components), yielding new algorithms for coastal waters. Furthermore, UV bands may promote the detection of harmful algal blooms such as red tides, because they produce UV absorbing compounds called mycosporine-like amino acids (Laurion et al. 2003). A mitigating factor is going to be whether or not the light instruments involved can be adequately calibrated in the UV domain. Most commercial sources of instrument calibration have a suitable capability in the visible part of the spectrum, but the capabilities in the UV portion remain largely unquantified.

#### 4.5 Primary Productivity

An important ecological property derived from oceanic remote sensing data is net primary production (NPP). This measurement is important, because it is a general indicator of the current health, and a monitor of future change, of an aquatic ecosystem. NPP can be defined as the amount of daily photosynthetically fixed carbon available to the first heterotrophic level (Lindeman 1942) or as gross photosynthesis minus diel respiration by the photosynthesizing organism (for the oceans, this is largely phytoplankton). Unfortunately, NPP cannot be directly measured from space, but it can be estimated from information on incident photosynthetically available radiation (PAR), phytoplankton biomass and its vertical distribution, mixed layer light levels, and the distribution of growth-limiting factors (e.g., micro- and macronutrients). Reducing uncertainties in NPP estimates requires validation of these required input variables, in addition to direct comparisons between measured and modeled NPP values for the water column.

Chlorophyll, or total pigment concentration, is an essential part of any NPP model. Uncertainties in remote sensing chlorophyll (pigment) products propagate through to NPP estimates and must be both minimized and well characterized. Presently, a primary source of uncertainty in chlorophyll (pigment) products derives from the inaccurate separation of pigment and CDOM absorption (Siegel et al. 2005). Approaches for addressing this issue and associated measurement requirements are discussed above in Sect. 4.4. In coastal or other optically complex waters, uncertainties in chlorophyll (pigment) retrievals can be significant and lead to large uncertainties in NPP estimates and carbon fluxes. In addition to using advanced products derived from spectral-matching algorithms (as opposed to band ratios), improved assessments of phytoplankton pigment absorption may be achieved through remote sensing measurements of Chl *a* fluorescence coupled with water-leaving radiance measurements in the long-wave ultraviolet

region (Huot et al. 2007). This approach requires an accurate assessment of incident PAR, fluorescence line heights, and fluorescence quantum yields. The latter two demand new and expanded field measurements, which will have to be properly validated and characterized.

Surface Chl *a* concentration is the traditional metric for phytoplankton biomass in NPP models. Chlorophyll (pigment) concentration, however, is not only a function of phytoplankton abundance, but is also strongly influenced by growth conditions within the mixed layer. This *physiological* variability is at the heart of major uncertainties in NPP products and remains a challenge to constrain. Historical approaches to this issue have involved the parameterization of physiological NPP model variables using relationships with specific physical properties—for example, temperature and incident PAR (Antoine et al. 1996, Behrenfeld and Falkowski 1997, and Behrenfeld et al. 2002)—or climatological descriptions of ecophysiological *provinces* derived from carbon fixation measurements using <sup>14</sup>C methods (Longhurst et al. 1995). These approaches, however, are highly empirical and considerable effort (including field and laboratory research) is needed to transition them to more mechanistic descriptions. If the transition can be made and applied to satellite data, then two important inquiries can be addressed: a) traditional behaviors in photosynthetic assimilation efficiencies, that is, carbon fixed per unit pigment biomass or absorption; and b) the properties whose mechanistic underpinnings are not yet fully understood, for example, covariations in light-limited and light-saturated pigment-normalized photosynthetic rates (Behrenfeld et al. 2004).

A different approach to ocean productivity modeling was recently proposed based on assessing phytoplankton carbon biomass using remote sensing particulate backscattering coefficients ( $b_{bp}$ ) from spectral-matching algorithms (Behrenfeld et al. 2005 and Westberry et al. 2007). The underlying concept of this approach is that  $b_{bp}$  covaries with particle abundance and, because of the relatively conserved nature of the particle size spectrum in natural waters, phytoplankton carbon. The advantage of the approach is that covariations in chlorophyll and carbon biomass reflect shifts in phytoplankton abundance, while their independent behavior tracks spatial and temporal variations in physiological status.

With this *carbon-based* approach, satellite observations provide information on both phytoplankton abundance and physiology. To derive NPP in this manner, variability in the physiological term—the carbon-to-chlorophyll ratio (C:Chl)—must be partitioned into a component attributable to photoacclimation (i.e., adjustments in C:Chl caused by changing light conditions) and a component ascribed to nutrient stress. The former requires an accurate description of incident PAR, mixed layer depths, and spectral attenuation coefficients, plus an improved understanding of community photoacclimation strategies and relationships between C:Chl growth rates. The carbon-based

approach will also benefit from the inclusion of new relationships accounting for light absorption by the full complement of photosynthetically-active pigments.

## 4.6 The Advanced Science Plan

The designated mission themes from the Advanced Science Plan, along with the corresponding high-priority research questions (Sect. 1), highlighted the science, technology, and mission concepts that the calibration and validation activity must help enable (Fig. 1). The associated research questions result in the identification of five mission and sensor scenarios in the following priority:

- An advanced oceanic radiometer (1 km spatial resolution) in low Earth orbit (LEO) to be enhanced with measurements of aerosol height and properties on a subsequent mission,
- A geostationary radiometer capable of surveying coastal regions at an improved spatial resolution (250 m or less) several times a day,
- A LEO high-spatial resolution (e.g., approximately 25 m) radiometer for estuarine and nearshore studies at even higher resolution,
- A variable fluorescence lidar, and
- A capability to estimate global mixed layer depths.

The last scenario may not be a direct satellite observation, but could be a model product generated with the assimilation of global satellite observations of SST, wind speeds, downwelling irradiation, etc.

To support the mission themes and assist in answering the research questions, the calibration and validation activity will contribute the following to the ensuing scientific investigations: a) provide continuous and consistent high-quality satellite-derived radiometric and biogeochemical data products; b) establish protocols—including levels of uncertainties and performance metrics—consistent with the agreed-upon requirements for *in situ* and laboratory measurements; c) engage the broader scientific community to participate in calibration and validation activities (field programs, round robins, workshops, etc.); d) provide AOP, IOP, and biogeochemical observations satisfying the quality requirements for calibration and validation activities; e) enable the development of new hardware and software, as well as its subsequent evaluation; and f) evaluate calibration and validation requirements for new campaigns (e.g., NACP, SOCP, etc.), missions (e.g., hyperspectral and UV radiometry capabilities), and measurements (e.g., physiology and functional types).

The implementation timeline for the satellite missions required to support the Advanced Science Plan involves three temporal horizons. The *immediate* (next 1–5 y) time frame will rely on current (SeaWiFS and MODIS) and operational (VIIRS) ocean color missions, which will be providing water-leaving radiances at the standard limited set of wavebands to support the continuation of the historical

chlorophyll record. None of these sensors satisfy the observational requirements to answer the four science questions posed above, and the recommendation is to construct and launch a hyperspectral radiometer with global imaging capabilities within this time period, or as soon as technologically possible. The expectation for the *near-term* (5–10 y) is the exploitation of the aforementioned advanced hyperspectral imager in conjunction with VIIRS and heritage (climatic) data. Towards the end of this period, the launch of a second-generation hyperspectral imager (with a scanning polarimeter and aerosol lidar) is anticipated, which would build upon the successes of the first hyperspectral sensor. The *long-term* (10–25 y) assessment assumes an advanced hyperspectral imager will have been flown successfully for at least 10 y, after which, the technology would be ready for transition to an operational phase (to replace the prior generation of fixed-wavelength instruments).

In terms of vicarious calibration, the schedule and requirements for the Advanced Science Plan do not conflict with what is being done right now. In fact, the highest-priority mission—the hyperspectral imager—builds on SeaWiFS and MODIS achievements. If changes are made to how vicarious calibration data are collected, the ability to support a hyperspectral mission with the selected alternative will have to be assessed. Because ocean color measurements are not fully independent, from a spectral point of view, a reduced dimensional (spectral) representation might be viable if it provides significant resource savings or reduced uncertainties. The future missions will unequivocally require new work, but this does not have to occur immediately: *the new calibration and validation activities can be phased in with the execution of the corresponding science plan components.*

Product validation will continue in much the same way as it has in the past, i.e., PI-supported research with independent field observations for comparison. SeaBASS will be continued and expanded to support this work. It is expected, however, that the suite of products will grow substantially from the current set of archive products, and this will require a more coordinated effort to best utilize the available ship time and to ensure observations are collected over the broadest possible range of bio-optical provinces. One of the recurring limitations of the past programs (SeaWiFS, MODIS, and SIMBIOS) was the inability to get complete data sets for algorithm development (e.g., AOPs, IOPs, phytoplankton pigments, CDOM, etc.) that were collected contemporaneously. As a result, uncertainties remain in terms of algorithm performance (e.g., chlorophyll *a*). In addition, the Advanced Science Plan outlines objectives, such as hazards and habitats, that have not been a focus in the past. This will require a new set of data product specifications, measurement requirements and strategies, protocols, and algorithms. Similarly, depending on the nature of the fluorescence lidar and its derived products, new strategies for validation sampling and measurements will probably be needed.

## 5. Strategic Summary

The primary objective of this planning document is to provide the organizational framework for addressing the long-term need to calibrate and validate oceanic biological and biogeochemical satellite data. The philosophy adopted here is to accept this as a practical problem that can be addressed with pragmatic solutions based primarily on the experience gained from prior satellite missions. A horizontal organizational scheme is anticipated wherein the overall leadership comes from the OBB Program Manager and a Calibration and Validation Team. The entire enterprise is split into two components of equal stature: a) calibration and validation, and b) satellite data processing. The respective success of these primary functions depend primarily (but not exclusively) on the radiometric accuracy of the on-orbit and *in situ* measurements, as well as the availability of the near-real time and archived (reprocessed) product suite.

The desire to create a *level playing field* extends into the disciplines required to make calibration and validation activities successful. In the planning for SeaWiFS and MODIS, the calibration and validation paradigm (Hooker and McClain 2000) emphasized the optical measurements with the biological and biogeochemical contributions relegated to secondary or tertiary importance (with the exception of chlorophyll *a*). Although such a philosophy might have been scientifically justified by the emphasis on the open ocean for those missions, the acknowledged future of ocean color includes a much wider array of significantly more complex problems both in the open and coastal ocean (Sect. 4.6). In fact, for many of the anticipated research areas significant contributions are expected from the modeling community, which was only represented in the original paradigm as part of the atmospheric correction algorithm. Consequently, the approach here is to ensure all of the needed disciplines are included at the same priority level and with no preconceived notions of resource allocations.

All of the plan's elements are seen to be interdependent, that is, they connect together like puzzle pieces, and when they are properly joined, a comprehensive capability for the entire activity emerges. For example, the calibration and validation paradigm is completely integrated with the satellite data processing capability. The full extent of the resulting competency depends very much on the resources available, but a much smaller *internal* core functionality provides enough of the total effort that significant progress can be made even during reduced budget cycles.

The scalability of the plan—that is, the scope of individual topics, the range and diversity of research objectives, and the number of investigators involved—is primarily accomplished through *external* connecting-core elements. These elements can be a combination of contractual agreements and peer-reviewed competitions, which are explicitly tied to the Carbon Cycle and Ecosystems Focus Area. The internal core scientists are expected to have

a breadth of skills and knowledge that extends throughout the connecting-core elements to ensure a strong union between the scientific community and the plan's principal elements, but the size of the group is not expected to expand or decrease appreciably over time (unless unforeseen significant changes in the broader program occur).

The detailed elements of the activity are based on the basic tasks of the two main components plus the current objectives of the OBB Program. The core elements for the calibration and validation component include a) publish protocols and performance metrics to ensure all data are to within community-approved specifications; b) verify uncertainty budgets and use the results to update a total uncertainty budget; c) manage the development and evaluation of instrumentation needed to satisfy the present and scheduled research goals of the overall activity; and d) coordinate international partnerships to provide the greatest diversity of global data, skills, and ideas. The core elements for the satellite data processing component are e) process and reprocess multisensor data; f) acquire, distribute, and archive data products; and g) implement new data products. Both main components have shared responsibilities for initializing and temporally monitoring satellite calibration with the most accurate methods available.

Connecting-core investigations include (but are not restricted to) atmospheric correction and characterization, standards and traceability, instrument and analysis round robins, field campaigns and vicarious calibration sites, *in situ* database, bio-optical algorithm (and product) validation, satellite characterization and vicarious calibration, and image processing software. An underlying objective within these activities is to ensure access to the highest-quality satellite and field data possible using sole-source capabilities where appropriate, e.g., an *in situ* database (SeaBASS), an image processing environment (SeaDAS), a central laboratory for pigment analysis, and a (recommended) capability for processing field data with a single processor (first for AOP and then for IOP data).

Public access to the data collected and used in calibration and validation activities is an important part of the plan. The philosophy advocated here is to make all data, including the calibration data (i.e., the satellite radiance and relevant *in situ* measurements) available to the community at large through the SeaBASS database, so those requiring unique processing procedures can implement them. The collection of high-quality data is an arduous task, however, the scientists producing such data need an appropriate amount of time to publish the scientific accomplishments associated with the data. Consequently, the entire community must agree and support a data policy for the collection and sharing of field observations.

The plan also includes an accountability process (which includes annual reviews evaluated by a team of experts and open to the wider community), creation of a Calibration and Validation Team (to help manage and oversee the activity), and a discussion of issues associated with the plan's

scientific focus. To ensure the latter has the broadest impact and that the scientists involved have a vested interest in the quality of the work being done, each element will have a set of scientific objectives. The latter must be endorsed by the wider community, and the work performed will be evaluated during the annual review. Core scientists are also expected to publish their achievements in technical reports and the peer-reviewed literature. The publication requirement is a distinguishing feature, because when results cannot get published or if no attempt is made to do so, *it means the scientific process has failed, and the community is left with more questions than answers, which is not an acceptable outcome.*

The OBB Program has traditionally focused on water-leaving radiances and chlorophyll *a* concentrations in the open ocean, as well as some experimental products (e.g., particulate inorganic carbon). The future product suite will be broadened, very likely move towards semi-analytical models (requiring IOPs plus other biogeochemical measurements), and place more emphasis on coastal (optically-complex) waters. This change in perspective will require advances in all aspects of the calibration and validation paradigm.

The challenges for satellite, field, and laboratory measurements will come from a diversity of disciplines and considerable preparation will be required to achieve success. The carbon-based modeling approach provides a good example of what is needed to improve a data product, in this case NPP: improved products from spectral-matching algorithms, advanced spaceborne sensors using considerably greater spectral resolution and range, characterized relationships between particulate backscattering and phytoplankton carbon and its sensitivity to shifts in particle size distribution, and assessed relationships between the spectral slope of particulate backscattering and the particle size distribution (Loisel et al. 2006). Improvements will also be needed in characterizing mixed layer light conditions (including the remote sensing of physiological mixing depths), as well as relating nutrient-dependent changes in carbon-to-chlorophyll ratios to growth rates and how these relationships differ between different types of nutrient stress (e.g., nitrogen limitation versus iron limitation).

The generation of consistent, high-quality data from new—and most likely more sophisticated—spaceborne instruments, multiple ocean color satellites, or a mix of these plus other types of satellite sensors, all require a diverse suite of very focused and closely coordinated activities ranging from the design and characterization of satellite instruments to the collection of field and laboratory data sets. There are many lessons learned from the SeaWiFS, MODIS, and SIMBIOS experiences. One of the most important is simply *the calibration and validation program requires a core group working full time on coordinating the activities and resolving the technical issues associated with data quality.*

Although the internal core group of scientists are expected to participate in many aspects of calibration and validation, the connecting-core researchers, the Calibration and Validation Team, and the wider scientific community are expected to help resolve a large diversity of problems and provide a community-wide approval of the procedures, priorities, and resources applied. Some of the most important issues are how best to a) satisfy the requirement for vicarious calibration sites, b) properly address optical complexity in coastal waters, c) develop bio-optical algorithms, and d) select the remote sensing data products. The SeaWiFS and MODIS experiences have demonstrated the need for multiple vicarious calibration sites, but the technology to be applied and the geographic distribution have yet to be determined. All of these issues could be resolved in parallel with the optical complexity issue if a common—but unequivocally suitable—measurement technique is chosen for all these issues. If a common (or centralized) deployment platform and instrument suite can be agreed upon, there is a chance for significant savings, both in terms of cost and overall uncertainties.

Anticipating unknown problems is always difficult, but building a flexible structure that can rapidly adapt to the unexpected is one of the best ways to be prepared. The aforementioned scalability of the plan proposed here is part of the philosophy of building a flexible capability. A real-world example of an unforeseen problem requiring a flexible response might be the need for a high-latitude, temporary vicarious calibration site for polarization studies. A low-cost, easily deployed capability—rather than a fixed site with a mostly static infrastructure—would provide many solution scenarios for unexpected problems (as long as the data are within the requisite performance requirements).

One of the challenges with an approach based on multiple vicarious calibration sites, whether they are temporary or not, is properly reconciling the differences in the satellite gains derived from each site. Assuming the *in situ* measurement capability at all sites are equally competent, the environmental conditions (e.g., aerosols, wind speed, sun glint contamination, etc.) will vary and produce a variance in the satellite gains. Quality control procedures will have to be developed (similar to those already being used, but expanded to accommodate the larger range in the variables) to ensure only the data from the best environmental conditions are used.

An excellent example of a centralized approach using multiple sites has already been demonstrated with the increasingly diverse deployment of SeaPRISM units. This above-water radiometry capability has matured sufficiently to satisfy the accuracy requirements for product validation while also providing excellent measurements in optically complex waters throughout the world ocean with almost negligible instrument fouling (Zibordi et al. 2006). Other advantages of this approach are it is based on commercial instrumentation, and the infrastructure to calibrate

the sensors, deploy the instruments (especially in coastal sites), and process the data already exists (AERONET).

A compelling inquiry, therefore, is determining what enhancements (e.g., sensor characterizations) are needed to certify a commercial device like SeaPRISM is acceptable for vicarious calibration exercises. A significant advantage of a network approach over a single site is how quickly the observations for a reliable gain factor analysis can be accrued. It took about three years of MOBY data, for example, to obtain the necessary 40 data points (Franz et al. 2007). Regardless of what solution is pursued, COTS hardware should be considered, because it provides a lower-cost and more flexible alternative to one-of-a-kind and hard-to-replicate instrumentation. The BOUSSOLE buoy, for example, is providing very good vicarious calibration data using low-cost and easily replicated radiometers.

Perhaps most importantly, both SeaPRISM and BOUSSOLE achieved a high level of success without using hyperspectral sensors. This shows the most difficult aspect of the vicarious calibration problem is most likely the (not-so-simple) requirement of making high quality observations in the harsh marine environment—that is, *measurements with a documented uncertainty in keeping with established performance metrics*—rather than puzzling out the intricacies of spectral response functions in spaceborne and *in situ* radiometers. This discussion of more practical alternatives is not meant to minimize the significance of understanding the spectral characteristics of the instruments involved; it is instead a candid affirmation that *the fight is ultimately in the field* and well beyond the nonetheless important work done in the laboratory. As the spectral diversity of remote sensing spreads into the UV and NIR or across many more channels, the ensuing complexity might require an alternative point of view, but for the fixed-wavelength satellites currently in operation or planned for launch (e.g., VIIRS), a rather basic and low-cost approach to vicarious calibration appears tenable and worth investigating.

Much has been learned since the early 1990s, and the plan presented here takes advantage of this knowledge to propel the entire calibration and validation enterprise into new areas of endeavor.

This document is based on providing a pragmatic solution to a practical problem, with the pragmatism coming from the lessons learned during the execution of historical and current (U.S.) ocean color satellite missions. It is worth noting in conclusion that an alternative—much more distributed—approach for calibrating and validating satellite-based ocean biology and biogeochemical measurements was tried during the EOS era, but it failed to deliver in many significant aspects of the overall enterprise. Consequently, the EOS approach is being revisited, which may lead to its replacement with the collocated architecture espoused in this plan—and *the tightly integrated approach has already demonstrated a capability of delivering at the required level of accuracy and timeliness*.

## ACKNOWLEDGMENTS

A number of individuals, groups, and organizations made significant contributions to NASA ocean color research that were critical in establishing the foundation of accomplishments—that is, the lessons learned—used in the plan presented here. For the authors, this legacy begins with Orbital Sciences Corporation (OSC), Santa Barbara Research Center, and ORBIMAGE who built, launched, and continue to maintain SeaWiFS: a spacecraft and instrument that has worked exceptionally well. It is important to remember, however, that several different GSFC engineering groups provided critical assistance to OSC in diagnosing problems with the spacecraft, launch vehicle, and instrument prior to launch.

During the prelaunch phase, NASA senior managers, particularly Bill Townsend, Dixon Butler, and Stan Snyder, were steadfast in their support. In addition, Bob Kirk and Mary Cleave oversaw the early phases through the acceptance of the SeaWiFS instrument and guided much of the spacecraft acceptance and launch, respectively. Throughout the process of getting the mission approved and launched, several Ocean Biogeochemistry Program Managers assisted and represented the Project at NASA headquarters (Marlon Lewis, Greg Mitchell, Robert Frouin, Jim Yoder, Janet Campbell, John Marra, Chuck Trees, and Paula Bontempi).

A number of international collaborations were—and continue to be—sources of key accomplishments, such as the Plymouth Marine Laboratory in the United Kingdom (Jim Aiken), the Joint Research Center in Italy (Giuseppe Zibordi), and the *Laboratoire d’Océanographie de Villefranche* in France (David Antoine, Hervé Claustre, and André Morel). Several members of the MODIS oceans team made a variety of major contributions including Wayne Esaias, Dennis Clark, Bob Evans, Howard Gordon, and Ken Carder.

More recently, the MCST has provided essential support in solving MODIS-A problems and will be an important partner in resolving MODIS-T data quality issues. The MCST has been led by Jack Xiong under the sensor PI, Vince Salomonson. The collaboration with the MCST has also led to a close working relationship with the analog group for VIIRS, the NPOESS Preparatory Project Instrument Characterization Support Team, also under Jack Xiong’s leadership. These collaborations have led to valuable insights into both the MODIS and VIIRS sensors as well as characterization methodologies and data. The collaborations with NIST (primarily Carol Johnson) were invaluable to the calibration of not only the SeaWiFS instrument, but also the *in situ* data.

The SIMBIOS program made many contributions to NASA ocean color activities, especially in the area of international cooperation and an international science team. The project worked extensively on the MOS, OCTS, POLDER, KOMPSAT, and MODIS-T data sets. Jim Mueller worked at GSFC for one year as an IPA to assist in getting the project started, and Giulietta Fargion is thanked for her invaluable participation in project management.

Throughout the time period of all these endeavors, Science Applications International Corporation and its subcontractors (Science Systems and Applications, Inc., and Futuretech Corporation) provided, and continue to provide, a talented and dedicated support staff. The ocean color research community has given unwavering support and guidance to the various satellite missions, which has been a great encouragement.

Finally, the broader scientific community is recognized for sharing their ideas and insights in a public review of the first draft of this document. The scientists who took time from their precious research activities or government duties to carefully ponder the complexities of such a far-reaching plan and make useful comments are gratefully acknowledged. The individuals who were approached after the review are thanked for their willingness to provide comments or material to address specific aspects of the document (notably Mike Behrenfeld).

## GLOSSARY

AAOT	<i>Acqua Alta</i> Oceanographic Tower
AERONET	Aerosol Robotic Network
AMT	Atlantic Meridional Transect
AOPs	Apparent Optical Properties
BOUSSOLE	<i>Bouée pour l’acquisition de Séries Optiques à Long Terme</i> (buoy for the acquisition of a long-term optical series).
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CDOM	Colored Dissolved Organic Matter
CDR	Climate Data Record
Chl <i>a</i>	Chlorophyll <i>a</i>
CHORS	Center for Hydro-Optics and Remote Sensing
CIMEL	Not an acronym, but the name of a sun photometer manufacturer.
COCOA	Coastal Ocean Carbon Observations and Applications
COTS	Commercial Off-the-Shelf
CZCS	Coastal Zone Color Scanner
DARR	Data Analysis Round-Robin
DARR-94	The first DARR activity (1994).
DARR-00	The second DARR activity (2000).
DOC	Dissolved Organic Carbon
EOS	Earth Observing System
GOCECP	Global Ocean Carbon, Ecosystems, and Coastal Processes
HPL	Horn Point Laboratory
HPLC	High Performance Liquid Chromatography
IOPs	Inherent Optical Properties
IPA	Intergovernmental Personnel Act
JGOFS	Joint Global Ocean Flux Study
JRC	Joint Research Centre
KOMPSAT	Korea Multi-Purpose Satellite
LEO	Low Earth Orbit
LOQ	Limit of Quantitation
LOV	<i>Laboratoire d’Océanographie de Villefranche</i> (Oceanographic Laboratory of Villefranche)
MCST	MODIS Characterization Support Team
MERIS	Medium Resolution Imaging Spectrometer
MOBY	Marine Optical Buoy
MODIS	Moderate Resolution Imaging Spectroradiometer
MODIS-A	MODIS on the Aqua spacecraft.
MODIS-T	MODIS on the Terra spacecraft.
MOS	Modular Optoelectronic Scanner

NACP	North Atlantic Carbon Program
NASA	National Aeronautics and Space Administration
NET	NIMBUS-7 Experiment Team
NIMBUS	Not an acronym, but a series of NASA experimental weather satellites containing a wide variety of atmosphere, ice, and ocean sensors.
NIR	Near-Infrared
NIST	National Institute of Standards and Technology
NPOESS	National Polar Orbiting Environmental Satellite System
NPP	Net Primary Production
OBB	Ocean Biology and Biogeochemistry
OCEaNS	Ocean Carbon, Ecosystem, and Near-Shore
OCTS	Ocean Color Temperature Scanner
OSC	Orbital Sciences Corporation
PAR	Photosynthetically Available Radiation
Perid	Peridinin
PFTs	Physiology and Functional Types
PI	Principal Investigator
POC	Particulate Organic Carbon
POLDER	Polarization Detecting Environmental Radiometer
PPIG	Primary Pigment
QA	Quality Assurance
RPD	Relative Percent Difference
RVS	Response Versus Scan
SeaBASS	SeaWiFS Bio-Optical Archive and Storage System
SeaDAS	SeaWiFS Data Analysis System
SeaHARRE	SeaWiFS HPLC Analysis Round-Robin Experiment
SeaHARRE-1	The first SeaHARRE activity (1999).
SeaHARRE-2	The second SeaHARRE activity (2002).
SeaHARRE-3	The third SeaHARRE activity (2004).
SeaPRISM	SeaWiFS Photometer Revision for Incident Surface Measurements
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SOCP	Southern Ocean Carbon Program
SIMBIOS	Sensor Intercomparison and Merger for Biological and Interdisciplinary Oceanic Studies
SIMRIC	SIMBIOS Radiometric Intercomparison
SIRREX	SeaWiFS Intercalibration Round-Robin Experiment
TChl <i>a</i>	Total chlorophyll <i>a</i>
UV	Ultraviolet
VIIRS	Visible and Infrared Imaging Radiometer Suite

## SYMBOLS

$A'$	Quality-assured subset of four SeaHARRE-2 methods.
$b_{bp}$	Particulate backscattering coefficient.
$d$	Sequential day of the year.
$E_d(0^+, \lambda)$	Spectral downward irradiance measured just above the sea surface (the global solar irradiance).

$F_0(\lambda, d)$	Mean extraterrestrial solar irradiance corrected for the Earth–Sun distance.
$L_u$	Upwelled radiance.
$L_W(\lambda)$	Spectral water-leaving radiance.
$[L_W(\lambda)]_N$	Spectral normalized water-leaving radiance.
$R_{rs}(\lambda)$	Spectral remote sensing reflectance.
$R_s$	Minimum resolution (between two pigments).
[TChl <i>a</i> ]	Total chlorophyll <i>a</i> concentration.
$X$	Independent observation.
$Y$	Dependent (reference) value.
$\lambda$	Wavelength (of light).
$\bar{\xi}$	Average precision.
$\bar{\xi}_{cal}$	Precision of the (calibration) dilution devices.
$\bar{\xi}_{inj}$	Average injection precision.
$\bar{\xi}_{t_R}$	Average retention time precision.
$\rho_w$	Radiance reflectance.
$ \bar{\psi} $	Average accuracy (based on the average absolute percent difference).
$ \bar{\psi} _{res}$	Average absolute percent differences of the residuals to the calibration fit for Chl <i>a</i> .

## REFERENCES

- Abbott, M.R., O.B. Brown, H.R. Gordon, K.L. Carder, R.E. Evans, F.E. Müller-Karger, and W.E. Esaias, 1994: Ocean Color in the 21<sup>st</sup> Century: A Strategy for a 20-Year Time Series. *NASA Tech. Memo. 104566, Vol. 17*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 20 pp.
- Aiken, J., N. Rees, S. Hooker, P. Holligan, A. Bale, D. Robins, G. Moore, R. Harris, and D. Pilgrim, 2000: The Atlantic Meridional Transect: Overview and synthesis of data. *Prog. Oceanogr.*, **45**, 257–312.
- Antoine, D., J-M. André, and A. Morel, 1996: Oceanic primary production 2. Estimation at global scale from satellite (coastal zone color scanner) chlorophyll. *Global Biogeochem. Cycles*, **10**, 57–69.
- , M. Chami, H. Claustre, A. Morel, G. Bécu, B. Gentili, F. Louis, J. Ras, E. Roussier, A.J. Scott, D. Tailliez, S.B. Hooker, P. Guevel, J-F. Desté, C. Dempsey, and D. Adams, 2006: BOUSSOLE: A Joint ESA, CNES, CNRS, and NASA Ocean Color Calibration and Validation Activity. *NASA Tech. Memo. 2006-214147*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 59 pp.
- Bailey, S.W., W.D. Robinson, and B.A. Franz, 2003: “Modifications to SeaWiFS NIR correction.” In: Patt, F.S., R.A. Barnes, R.E. Eplee, Jr., B.A. Franz, W.D. Robinson, G.C. Feldman, S.W. Bailey, J. Gales, P.J. Werdell, M. Wang, R. Frouin, R.P. Stumpf, R.A. Arnone, R.W. Gould, Jr., P.M. Martinolich, V. Ransibrahmanakul, J.E. O’Reilly, and J.A. Yoder, 2003: Algorithm Updates for the Fourth SeaWiFS Data Reprocessing, *NASA Tech. Memo. 2003-206892, Vol. 22*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 26–28.

- Baith, K., R. Lindsay, G. Fu, and C.R. McClain, 2001: Data analysis system developed for ocean color satellite sensors. *Eos, Trans., Amer. Geophys. Union*, **82**, 202.
- Barnes, R.A., W.L. Barnes, W.E. Esaias, and C.R. McClain, 1994a: Prelaunch Acceptance Report for the SeaWiFS Radiometer. *NASA Tech. Memo. 104566, Vol. 22*, S.B. Hooker, E.R. Firestone, and J.G. Acker, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 32 pp.
- , A.W. Holmes, W.L. Barnes, W.E. Esaias, C.R. McClain, and T. Svitek, 1994b: SeaWiFS Prelaunch Radiometric Calibration and Spectral Characterization. *NASA Tech. Memo. 104566, Vol. 23*, S.B. Hooker, E.R. Firestone, and J.G. Acker, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 55 pp.
- , A.W. Holmes, and W.E. Esaias, 1995: Stray Light in the SeaWiFS Radiometer. *NASA Tech. Memo. 104566, Vol. 31*, S.B. Hooker, E.R. Firestone, and J.G. Acker, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 76 pp.
- , R.E. Eplee, Jr., G.M. Schmidt, F.S. Patt, and C.R. McClain, 2001: The calibration of SeaWiFS, part 1: Direct techniques, *Appl. Opt.*, **40**, 6,682–6,700.
- Behrenfeld, M.J., and P.G. Falkowski, 1997: Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.*, **42**, 1–20.
- , E. Marañón, D.A. Siegel, and S.B. Hooker, 2002: A photoacclimation and nutrient based model of light-saturated photosynthesis for quantifying oceanic primary production. *Mar. Ecol. Prog. Ser.*, **228**, 103–117.
- , O. Prasil, M. Babin, and F. Bruyant, 2004: In search of a physiological basis for covariations in light-limited and light-saturated photosynthesis. *J. Phycol.*, **40**, 4–25.
- , E. Boss, D.A. Siegel, and D.M. Shea, 2005: Carbon-based ocean productivity and phytoplankton physiology from space. *Global Biogeochem. Cycles*, **19**, GB1006, doi: 10.1029/2004GB002299.
- Chang, G., K. Mahoney, A. Briggs-Whitmire, D.D.R. Kohler, C.D. Mobley, M. Lewis, M. Moline, E. Boss, M. Kim, W. Philpot, and T.D. Dickey, 2004: The new age of hyperspectral oceanography. *Oceanogr.*, **17**, 16–23.
- Clark, D., H.R. Gordon, K.J. Voss, Y. Ge, W. Broenkow, and C. Trees, 1997: Validation of atmospheric correction over the oceans. *J. Geophys. Res.*, **102**, 17,209–17,217.
- Claustre, H., S.B. Hooker, L. Van Heukelem, J-F. Berthon, R. Barlow, J. Ras, H. Sessions, C. Targa, C. Thomas, D. van der Linde, and J-C. Marty, 2004: An intercomparison of HPLC phytoplankton pigment methods using *in situ* samples: Application to remote sensing and database activities. *Mar. Chem.*, **85**, 41–61.
- Eplee, R.E., Jr., W.D. Robinson, S.W. Bailey, D.K. Clark, P.J. Werdell, M. Wang, R.A. Barnes, and C.R. McClain, 2001: The calibration of SeaWiFS, part 2: Vicarious techniques, *Appl. Opt.*, **40**, 6,701–6,718.
- Esaias, W.E., M.R. Abbott, I. Barton, O.B. Brown, J.W. Campbell, K.L. Carder, D.K. Clark, R.H. Evans, F.E. Hoge, H.R. Gordon, W.M. Balch, R. Letelier, and P.J. Minnett, 1998: An overview of MODIS capabilities for ocean science observations. *IEEE Trans. Geosci. Remote Sens.*, **36**, 1,250–1,265.
- Evans, R.H., and H.R. Gordon, 1994: CZCS system calibration: A retrospective examination. *J. Geophys. Res.*, **99**, 7,293–7,307.
- Firestone, E.R., and S.B. Hooker, 2004: SeaWiFS Postlaunch Technical Report Series Final Cumulative Index. *NASA Tech. Memo. 2004–206892, Vol. 29*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 44 pp.
- Franz, B.A., P.J. Werdell, G. Meister, S.W. Bailey, R.E. Eplee, Jr., G.C. Feldman, E. Kwiatkowska, C.R. McClain, F.S. Patt, and D. Thomas, 2005: The continuity of ocean color measurements from SeaWiFS to MODIS. *Proc. SPIE, Earth Observing Systems X*, San Diego, California.
- , S.W. Bailey, P.J. Werdell, C.R. McClain, 2007: Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry. *Appl. Opt.*, (in press).
- Fu, G., B.D. Scheiber, K.J. Settle, M. Darzi, C.R. McClain and K.R. Arrigo, 1996: SeaDAS: A processing package for ocean color satellite imagery, *Proc. Twelfth Int. Conf. on Interactive Inform. and Processing Systems for Meteorol., Oceanogr., and Hydrol.*, American Meteorological Society, Boston, Massachusetts, 451–456.
- Gordon, H.R., 1987: Calibration requirements and methodology for remote sensors viewing the ocean in the visible. *Remote Sens. Environ.*, **22**, 103–126.
- , D.K. Clark, J.L. Mueller, and W.A. Hovis, 1980: Phytoplankton pigments from the NIMBUS-7 Coastal Zone Color Scanner: Comparisons with surface measurements. *Science*, **210**, 63–66.
- , and —, 1981: Clear water radiances for atmospheric correction of coastal zone color scanner imagery. *Appl. Opt.*, **20**, 4,175–4,180.
- , —, J.W. Brown, O.B. Brown, R.H. Evans, and W.W. Broenkow, 1983: Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison of ship determinations and CZCS estimates. *Appl. Opt.*, **22**, 20–36.
- Holben, B.N., T.F. Eck, I. Slutsker, D. Tanré, J.P. Buis, A. Setzer, E. Vermote, J.A. Reagan, Y.J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov, 1998: AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.*, **66**, 1–16.
- Hooker, S.B., and W.E. Esaias, 1993: An overview of the SeaWiFS project. *Eos, Trans., Amer. Geophys. Union*, **74**, 241–246.
- , C.R. McClain, J.K. Firestone, T.L. Westphal, E-n. Yeh, and Y. Ge, 1994: The SeaWiFS Bio-Optical Archive and Storage System (SeaBASS), Part 1. *NASA Tech. Memo. 104566, Vol. 20*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 40 pp.

- , and S. Maritorena, 2000: An evaluation of oceanographic radiometers and deployment methodologies. *J. Atmos. Ocean. Technol.*, **17**, 811–830.
- , and C.R. McClain, 2000: The calibration and validation of SeaWiFS data. *Prog. Oceanogr.*, **45**, 427–465.
- , H. Claustre, J. Ras, L. Van Heukelem, J-F. Berthon, C. Targa, D. van der Linde, R. Barlow, and H. Sessions, 2000a: The First SeaWiFS HPLC Analysis Round-Robin Experiment (SeaHARRE-1). *NASA Tech. Memo. 2000–206892, Vol. 14*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 42 pp.
- , G. Zibordi, J-F. Berthon, S.W. Bailey, and C.M. Pietras, 2000b: The SeaWiFS Photometer Revision for Incident Surface Measurement (SeaPRISM) Field Commissioning. *NASA Tech. Memo. 2000–206892, Vol. 13*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 24 pp.
- , G. Zibordi, J-F. Berthon, D. D’Alimonte, S. Maritorena, S. McLean, and J. Sildam, 2001: Results of the Second SeaWiFS Data Analysis Round Robin, March 2000 (DARR-00). *NASA Tech. Memo. 2001–206892, Vol. 15*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 71 pp.
- , S. McLean, J. Sherman, M. Small, G. Lazin, G. Zibordi, and J.W. Brown, 2002: The Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7), March 1999. *NASA Tech. Memo. 2002–206892, Vol. 17*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 69 pp.
- , G. Zibordi, J-F. Berthon, and J.W. Brown, 2004: Above-water radiometry in shallow, coastal waters. *Appl. Opt.*, **43**, 4,254–4,268.
- , and G. Zibordi, 2005: Platform perturbations in above-water radiometry. *Appl. Opt.*, **44**, 553–567.
- , L. Van Heukelem, C.S. Thomas, H. Claustre, J. Ras, L. Schlüter, J. Perl, C. Trees, V. Stuart, E. Head, R. Barlow, H. Sessions, L. Clementson, J. Fishwick, C. Llewellyn, and J. Aiken, 2005: The Second SeaWiFS HPLC Analysis Round-Robin Experiment (SeaHARRE-2). *NASA Tech. Memo. 2005–212785*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 112 pp.
- Huot, Y., C.A. Brown, and J.J. Cullen, 2007: Simultaneous inversion of reflectance, the diffuse attenuation coefficient, and sun-induced fluorescence in coastal waters. *J. Geophys. Res.*, (in press).
- Johnson, B.C., S.S. Bruce, E.A. Early, J.M. Houston, T.R. O’Brian, A. Thompson, S.B. Hooker, and J.L. Mueller, 1996: The Fourth SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-4), May 1995. *NASA Tech. Memo. 104566, Vol. 37*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 65 pp.
- , E.A. Early, R.E. Eplee, Jr., R.A. Barnes, and R.T. Cafrey, 1999a: The 1997 Prelaunch Radiometric Calibration of SeaWiFS. *NASA Tech. Memo. 1999–206892, Vol. 4*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 51 pp.
- , H.W. Yoon, S.S. Bruce, P-S. Shaw, A. Thompson, S.B. Hooker, R.E. Eplee, Jr., R.A. Barnes, S. Maritorena, and J.L. Mueller, 1999b: The Fifth SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-5), July 1996. *NASA Tech. Memo. 1999–206892, Vol. 7*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, 75 pp.
- Joint Global Ocean Flux Study, 1991: JGOFS Core Measurements Protocols. *JGOFS Report No. 6*, Scientific Committee on Oceanic Research, 40 pp.
- Laurion, I., F. Blouin, and S. Roy, 2003: The quantitative filter technique for measuring phytoplankton absorption: Interference by MAAs in the UV waveband. *Limnol. Oceanogr. Meth.*, **1**, 1–9.
- Lee, Z., and K.L. Carder, 2004: Absorption spectrum of phytoplankton pigments derived from hyperspectral remote-sensing reflectance. *Remote Sens. Environ.*, **89**, 361–368.
- Lindeman, R.L., 1942: The trophic-dynamic aspect of ecology. *Ecology*, **23**, 399–418.
- Loisel, H., J-M. Nicolas, A. Sciandra, D. Stramski, and A. Poteau, 2006: Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean. *J. Geophys. Res.*, **111**, C09024 doi:10.1029/2005JC003367.
- Longhurst, A., S. Sathyendranath, T. Platt, and C. Caverhill, 1995: An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.*, **17**, 1,245–1,271.
- McClain, C.R., W.E. Esaias, W. Barnes, B. Guenther, D. Endres, S.B. Hooker, B.G. Mitchell, and R. Barnes, 1992: SeaWiFS Calibration and Validation Plan. *NASA Tech. Memo. 104566, Vol. 3*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 41 pp.
- , and G.S. Fargion, 1999: SIMBIOS Project 1998 Annual Report. *NASA Tech. Memo. 1999–208645*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 105 pp.
- , W. Esaias, G. Feldman, R. Frouin, W. Gregg, and S. Hooker, 2002: The Proposal for the NASA Sensor Intercalibration and Merger for Biological and Interdisciplinary Oceanic Studies (SIMBIOS) Program, 1995. *NASA Tech. Memo. 2002–210008*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 54 pp. plus graphical plates.
- , G.C. Feldman, and S.B. Hooker, 2004: An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. *Deep Sea Res. II*, **51**, 5–42.

- , S. Hooker, G. Feldman, and P. Bontempi, 2006: Satellite data for ocean biology, biogeochemistry, and climate research. *Eos, Trans., Amer. Geophys. Union*, **87**, 337, 343.
- Meister, G., P. Abel, K. Carder, A. Chapin, J. Cooper, C. Davis, M. Godin, D. Goebel, G. Fargion, R. Frouin, D. Korwan, R. Maffione, C. McClain, S. McLean, D. Menzies, A. Poteau, J. Robertson, and J. Sherman, 2002: The First SIMBIOS Radiometric Intercomparison (SIMRIC-1), April–September 2001. *NASA Tech. Memo. 2002–210006*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 60 pp.
- , —, R. Barnes, J. Cooper, C. Davis, D. English, G. Fargion, M. Feinholz, R. Frouin, F. Hoge, D. Korwan, G. Lazin, C. McClain, S. McLean, D. Menzies, A. Poteau, J. Robertson, J. Sherman, K. Voss, and J. Yungel, 2003: The Second SIMBIOS Radiometric Intercomparison (SIMRIC-2), March–November 2002. *NASA Tech. Memo. 2002–210006, Vol. 2*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 65 pp.
- Mitchell, G., A. Bricaud, K. Carder, J. Cleveland, G. Ferrari, R. Gould, M. Kahru, M. Kishino, H. Maske, T. Moisan, L. Moore, N. Nelson, D. Phinney, R. Reynolds, H. Sosik, D. Stramski, S. Tassan, C.C. Trees, A. Weidemann, J. Wieland, and A. Vodacek, 2000: “Determination of spectral absorption coefficients of particles, dissolved material and phytoplankton for discrete water samples.” In: Fargion, G.S., and J.L. Mueller, Eds., *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 2. NASA Tech. Memo. 2000–209966/Rev2*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 125–153.
- , M. Kahru, J. Wieland, and M. Stramska, 2003: “Determination of spectral absorption coefficients of particles, dissolved material and phytoplankton for discrete water samples.” In: Mueller, J.L., G.S. Fargion, and C.R. McClain, Eds., *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume IV. NASA Tech. Memo. 2003–211621/Rev4–Vol.IV.*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 39–64.
- Morel, A., and L. Prieur, 1977: Analysis of variations in ocean color. *Limnol. Oceanogr.*, **22**, 709–722.
- Mueller, J.L., 1993: The First SeaWiFS Intercalibration Round-Robin Experiment, SIRREX-1, July 1992. *NASA Tech. Memo. 104566, Vol. 14*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 60 pp.
- , 1995: “Comparison of irradiance immersion coefficients for several marine environmental radiometers (MERs).” In: Mueller, J.L., R.S. Fraser, S.F. Biggar, K.J. Thome, P.N. Slater, A.W. Holmes, R.A. Barnes, C.T. Weir, D.A. Siegel, D.W. Menzies, A.F. Michaels, and G. Podesta, *Case Studies for SeaWiFS Calibration and Validation, Part 3. NASA Tech. Memo. 104566, Vol. 27*, S.B. Hooker, E.R. Firestone, and J.G. Acker, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 3–15.
- , 2000: “Overview of measurement and data analysis protocols.” In: Fargion, G.S., and J.L. Mueller, *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 2. NASA Tech. Memo. 2000–209966*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 87–97.
- , 2002: “Overview of measurement and data analysis protocols.” In: Mueller, J.L., and G.S. Fargion, *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 3, Vol. 1. NASA Tech. Memo. 2002–210004/Rev3–Vol1*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 123–137.
- , 2003: “Overview of measurement and data analysis methods.” In: Mueller, J.L., and 17 Coauthors, *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measurements and Data Analysis Protocols. NASA Tech. Memo. 2003–211621/Rev4–Vol.III*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 1–20.
- , and R.W. Austin, 1992: *Ocean Optics Protocols for SeaWiFS Validation. NASA Tech. Memo. 104566, Vol. 5*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 43 pp.
- , B.C. Johnson, C.L. Cromer, J.W. Cooper, J.T. McLean, S.B. Hooker, and T.L. Westphal, 1994: The Second SeaWiFS Intercalibration Round-Robin Experiment, SIRREX-2, June 1993. *NASA Tech. Memo. 104566, Vol. 16*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 121 pp.
- , and R.W. Austin, 1995: *Ocean Optics Protocols for SeaWiFS Validation, Revision 1. NASA Tech. Memo. 104566, Vol. 25*, S.B. Hooker, E.R. Firestone, and J.G. Acker, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 66 pp.
- , B.C. Johnson, C.L. Cromer, S.B. Hooker, J.T. McLean, and S.F. Biggar, 1996: The Third SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-3), 19–30 September 1994. *NASA Tech. Memo. 104566, Vol. 34*, S.B. Hooker, E.R. Firestone, and J.G. Acker, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 78 pp.
- , and A. Morel, 2003: “Fundamental Definitions, Relationships and Conventions.” In: Mueller, J.L., and 4 Coauthors, *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume I: Radiometric Measurements and Data Analysis Protocols. NASA Tech. Memo. 2003–211621/Rev4–Vol.I*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 11–30.
- O’Reilly, J.E., S. Maritorena, B.G. Mitchell, D.A. Siegel, K.L. Carder, S.A. Garver, M. Kahru, and C. McClain, 1998: Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.*, **103**, 24,937–24,953.
- Riley, T., and S. Bailey, 1998: The Sixth SeaWiFS/SIMBIOS Intercalibration Round-Robin Experiment (SIRREX-6) August–December 1997. *NASA Tech. Memo. 1998–206878*, NASA Goddard Space Flight Center, Greenbelt, Maryland, 26 pp.

- Siegel, D.A., M.C. O'Brien, J.C. Sorensen, D.A. Konnoff, E.A. Brody, J.L. Mueller, C.O. Davis, W.J. Rhea, and S.B. Hooker, 1995: Results of the SeaWiFS Data Analysis Round-Robin (DARR-94), July 1994. *NASA Tech. Memo. 104566, Vol. 26*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 58 pp.
- , M. Wang, S. Maritorena, and W. Robinson, 2000: Atmospheric correction of satellite ocean-color imagery: the black pixel assumption. *Appl. Optics*, **39**, 3582–3591.
- , S. Maritorena, N.B. Nelson, and M.J. Behrenfeld, 2005: Independence and interdependencies among global ocean color properties: Reassessing the bio-optical assumption. *J. Geophys. Res.*, **110**, C07011, doi:10.1029/2004JC002527.
- Van Heukelem, L., and C. Thomas, 2001: Computer-assisted high-performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments. *J. Chromatogr. A.*, **910**, 31–49.
- , —, and P.M. Glibert, 2002: Sources of Variability in Chlorophyll Analysis by Fluorometry and High-Performance Liquid Chromatography in a SIMBIOS Intercalibration Exercise. *NASA Tech. Memo. 2002-211606*, G.S. Fargion and C.R. McClain, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 50 pp.
- Werdell, P.J., S. Bailey, G. Fargion, C. Pietras, K. Knobel-spieße, G. Feldman, and C. McClain, 2003: Unique data repository facilitates ocean color satellite validation. *Eos, Trans., Amer. Geophys. Union*, **84**, 379–385.
- Westberry, T., M.J. Behrenfeld, D.A. Siegel, E. Boss, 2007: Carbon-based primary productivity modeling with vertically resolved photophysiology. *Global Biogeochem. Cycles*, (submitted).
- Zibordi, G., D. D'Alimonte, D. van der Linde, J-F. Berthon, S.B. Hooker, J.L. Mueller, G. Lazin, and S. McLean, 2002a: The Eighth SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-8), September–December 2001. *NASA Tech. Memo. 2002-206892, Vol. 21*, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 39 pp.
- , S.B. Hooker, J-F. Berthon, and D. D'Alimonte, 2002b: Autonomous above-water radiance measurements from an offshore platform: A field assessment experiment. *J. Atmos. Oceanic Technol.*, **19**, 808–819.
- , F. Mélin, S.B. Hooker, D. D'Alimonte, and B. Holben, 2004: An autonomous above-water system for the validation of ocean color radiance data. *Trans. IEEE Trans. Geosci. Remote Sens.*, **42**, 401–415.
- , B. Holben, S.B. Hooker, F. Mélin, J-F. Berthon, I. Slutsker, D. Giles, D. Vandemark, H. Feng, K. Rutledge, G. Schuster, and A. Al Mandoos, 2006: A network for standardized ocean color validation measurements. *Eos, Trans., Amer. Geophys. Union*, **84**, 293, 297.

**REPORT DOCUMENTATION PAGE**

*Form Approved  
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b> 03-07-2007		<b>2. REPORT TYPE</b> Special Publication		<b>3. DATES COVERED (From – To)</b>	
<b>4. TITLE AND SUBTITLE</b>  NASA Strategic Planning Document: A Comprehensive Plan for the Long-Term Calibration and Validation of Oceanic Biogeochemical Satellite Data				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b>	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b>  Stanford B. Hooker, Charles R. McClain, and Antonio Mannino				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>  Goddard Space Flight Center Greenbelt, MD 20771				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>  2007-00805-0	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  National Aeronautics and Space Administration Washington, DC 20546-0001				<b>10. SPONSORING/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSORING/MONITORING REPORT NUMBER</b>  SP-2007-214152	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b>  Unclassified-Unlimited, Subject Category: 48 Report available from the NASA Center for Aerospace Information, 7115 Standard Drive, Hanover, MD 21076. (301)621-0390					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b>  This document establishes a long-term capability for calibrating and validating oceanic biogeochemical satellite data, and involves two components of equal stature: calibration and validation plus satellite data processing. The sub-elements are based on the basic tasks of the two main components plus the current objectives of the Carbon Cycle and Ecosystems Roadmap. The former is distinguished by an internal core set of responsibilities and the latter is facilitated through an external connecting-core ring of competed or contracted activities. The core elements for calibration and validation include publish protocols and performance metrics; verify uncertainty budgets; manage the development and evaluation of instrumentation; and coordinate international partnerships. The core elements for satellite data processing are process and reprocess multisensor data; acquire, distribute, and archive data products; and implement new data products. Both components share responsibilities for initializing and temporally monitoring satellite calibration. Connecting-core elements involve atmospheric correction and characterization, standards and traceability, instrument and analysis round robins, field campaigns and vicarious calibration sites, <i>in situ</i> database, bio-optical algorithm validation, satellite characterization and vicarious calibration, and image processing software. The plan also includes an accountability process, creating a Calibration and Validation Team, and a discussion of issues associated with the plan's scientific focus.					
<b>15. SUBJECT TERMS</b>  Ocean Color, SeaWiFS, MODIS, Ocean Biogeochemistry, Calibration and Validation					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b>	<b>b. ABSTRACT</b>	<b>c. THIS PAGE</b>			Stanford B. Hooker
Unclassified	Unclassified	Unclassified	Unclassified	31	<b>19b. TELEPHONE NUMBER (Include area code)</b> 301-286-9503



